

附件 1

5G 高新视频—互动视频技术白皮书（2020）

国家广播电视总局科技司

2020 年 8 月

前言

当前，移动信息技术飞速发展，5G 技术已经成为国际通信科技巨头竞争的新焦点，世界各国纷纷将 5G 建设视为重要目标。5G 技术具有“大带宽、低时延、广连接”等特点，在 5G 众多应用场景中，视频被公认为是 5G 时代最重要和最早开展的业务，越来越受到社会各界的广泛关注。在 5G、超高清、虚拟现实等新兴技术催生下，广电行业视听内容的生产和传播即将发生新变革。国家广播电视总局顺应技术革命浪潮，抢抓 5G 发展机遇，深入推进 5G 条件下广播电视供给侧结构性改革，推动构建 5G 视频新业态，提出了“5G 高新视频”的概念。

5G 高新视频是指 5G 环境下具有“更高技术格式、更新应用场景、更美视听体验”的视频。其中，“高”是指视频融合 4K/8K、3D、VR/AR/MR、高帧率（HFR）、高动态范围（HDR）、广色域（WCG）等高新技术格式；“新”是指具有新奇的影像语言和视觉体验的创新应用场景，能够吸引观众兴趣并促使其产生消费。在 5G 环境下，广电行业将发挥在视音频内容创意、生产、制作和传播方面的丰富资源和天然优势，以文化创意为牵引，以科技创新为支撑，培育孵化 5G 高新视频的新技术、新应用、新产品、新模式、新业态和新消费，为观众带来更美的视听体验，更好地满足人民群众日益增长的精神文化新需求和新期待。

互动视频作为高新视频业态的重要组成部分，是指以“非线性视频”内容为主线，在“非线性视频”内容上开展的可支持时间域互动、空间域互动、事件型互动的内容互动视频业务，该业务具有分支剧情选择、视角切换、画面互动等交互能力，能够为用户带来强参与感、强沉浸度的互动观看体验。

为更好地推动互动视频发展，引导与规范互动视频行业应用，国家广播电视总局科技司组织广播电视规划院等广播电视行业内外相关机构、企业公司等开展了互动视频研究工作，编制了《5G 高新视频-互动视频技术白皮书（2020 版）》（以下简称白皮书）。

本白皮书分析了互动视频的业务背景和发展现状，梳理了互动视频应用场景，对互动视频关键技术和端到端解决方案进行了深入探讨，总结了国内媒体机构的互动视频成功实践案例，展望了互动视频发展前景。希望本白皮书能为各单位应用互动视频提供参考和借鉴，为推动广播电视行业转型升级注入新动能、激发新活力，促进广播电视和网络视听行业高质量创新性发展。

本白皮书指导单位：国家广播电视总局科技司

本白皮书主要起草单位：国家广播电视总局广播电视规划院、中央广播电视总台、北京爱奇艺科技有限公司、腾讯科技（深圳）有限公司、湖南快乐阳光互动娱乐传媒有限公司、智令互动（深圳）科技有限公司、优酷信息技术（北京）有限公司、华为技术有限公司。

本白皮书主要编写指导：孙苏川、关丽霞。

本白皮书主要起草人：邓向冬、刘文翰、梅剑平、孙可、杨光、刘里、薛澈、张瑞圣、房秀强、王建、宁金辉、黄卓伟、王涛、杨子斌、孟庆春、洪薇、苏衍、毕蕴瑾、侯亮、陈翌、王斐、覃千洛。

目 录

缩略语.....	4
1. 互动视频概述.....	6
1.1 互动视频定义.....	6
1.2 互动视频发展历程.....	8
1.3 互动视频国内外应用发展现状.....	9
1.3.1 国际互动视频应用发展现状.....	9
1.3.2 国内互动视频应用发展现状.....	10
1.4 互动视频应用场景.....	12
1.4.1 互动影视剧.....	12
1.4.2 互动综艺.....	12
1.4.3 互动短视频.....	12
1.4.4 互动影像游戏.....	12
1.4.5 文化旅游.....	13
1.4.6 体育竞技.....	14
1.4.7 在线教育.....	15
1.4.8 电商购物.....	15
1.4.9 互联网广告.....	16
1.5 互动视频应用机遇.....	16
1.6 互动视频应用挑战.....	17
2. 互动视频关键技术.....	18
2.1 互动视频整体技术框架.....	18
2.2 信令数据封装技术.....	18
2.3 多视点编码技术.....	20
2.4 智能编码技术.....	22
2.5 大带宽、低时延传输技术.....	22
2.6 网络切片技术.....	24
2.7 互动播放引擎技术.....	24
2.8 智能预测技术.....	25
2.9 5G 环境对互动视频关键技术的影响.....	25
3. 互动视频端到端解决方案.....	28
3.1 互动视频端到端系统架构.....	28
3.2 互动视频生产发布流程.....	28
3.3 互动视频创作流程.....	29
3.3.1 内容创作.....	29
3.3.2 现场拍摄.....	29
3.3.3 视频制作.....	30
3.4 互动视频制作平台.....	31
3.4.1 素材管理.....	31
3.4.2 故事线编辑.....	32
3.4.3 互动组件管理.....	32

3.4.4	互动触发条件.....	33
3.4.5	互动应用模板.....	33
3.4.6	互动视频预览.....	34
3.4.7	互动视频提交.....	34
3.5	互动视频服务平台.....	34
3.5.1	媒体资产管理.....	34
3.5.2	网络分发.....	36
3.5.3	用户账号管理.....	37
3.5.4	历史观看存档.....	37
3.5.5	数据采集分析.....	38
3.6	互动视频播放系统.....	40
3.6.1	播放器.....	40
3.6.2	互动播放引擎.....	40
3.6.3	互动方式.....	42
3.6.4	用户操作响应时延约束.....	42
3.7	跨平台兼容实现.....	43
3.7.1	互动协议跨平台解析标准化.....	44
3.7.2	互动播放跨平台兼容适配.....	46
4.	互动视频安全.....	49
4.1	内容安全.....	49
4.1.1	内容合规性安全审核.....	49
4.1.2	内容合理性安全审核.....	49
4.2	版权保护.....	49
4.2.1	互动视频制作过程版权保护.....	50
4.2.2	互动视频审核过程版权保护.....	50
4.2.3	互动视频分发过程版权保护.....	51
4.2.4	互动视频播放过程版权保护.....	51
4.3	介质安全.....	51
4.3.1	介质管理机制.....	51
4.3.2	灾备管理机制.....	52
4.4	运行环境安全.....	52
4.4.1	互动脚本安全.....	52
4.4.2	隔离与隐私保护.....	52
4.4.3	沙箱执行环境.....	52
4.5	平台与传输安全.....	53
4.5.1	平台安全.....	53
4.5.2	传输安全.....	53
4.6	用户隐私安全.....	53
4.6.1	用户隐私授权机制.....	53
4.6.2	用户数据防护机制.....	54
5.	互动视频应用展望.....	55
附录 A	互动影视剧案例.....	56
A.1	案例 1.....	56
A.2	案例 2.....	58

A.3 案例3.....	61
附录B 互动综艺案例.....	65
附录C 互动短视频案例.....	67
参考文献.....	69

缩略语

缩略语	英文全称	中文全称
4G	4th Generation mobile networks	第四代移动通信网络
5G	5th Generation mobile networks	第五代移动通信网络
ABU	Asia-Pacific Broadcasting Union	亚洲-太平洋广播联盟
AI	Artificial Intelligence	人工智能
API	Application Programming Interface	应用程序接口
APP	Application	应用软件
ATSC	Advanced Television Systems Committee	高级电视系统委员会
AVG	Adventure Game	冒险游戏
AVS2	The second generation of Audio Video coding Standard	第二代数字音视频编解码技术标准
CDN	Content Delivery Network	内容分发网络
DRM	Digital Rights Management	数字版权管理
ETSI	European Telecommunications Standards Institute	欧洲电信标准化协会
FIFO	First Input First Output	先入先出队列
FlexE	Flex Ethernet	灵活以太网
FOV	Field of View	视场角
HbbTV	Hybrid Broadcast Broadband Television	混合广播宽带电视
HTML5	Hyper Text Markup Language5.0	超文本标记语言 5.0
HTTPS	Hyper Text Transfer Protocol over SecureSocket	超文本传输安全协议
IBB	Integrated Broadcast-Broadband System	综合广播宽带系统
iOS	iPhone Operation System	苹果操作系统
IP 网络	Internet Protocol	网间互联协议
IP 版权	Intellectual Property	知识产权
ISO	International Organization for Standardization	国际标准化组织
ITU	International Telecommunication Union	国际电信联盟
MCN	Multi-Channel Network	多频道网络
MEC	Multi-access Edge Computing	多址边缘计算
MHP	Multimedia Home Platform	多媒体家庭平台
MPEG	Moving Picture Experts Group	动态图像专家组
PGC	Professional Generated Content	专业生产内容

SDK	Software Development Kit	软件开发工具包
SoC	System on a Chip	系统级芯片
TLS	Transport Layer Security	安全传输层
TS	Transport Stream	传送流
UGC	User Generated Content	用户生产内容
UPS	Uninterruptible Power System	不间断电源系统
VR	Virtual Reality	虚拟现实
WLAN	Wireless Local Area Network	无线局域网

1. 互动视频概述

1.1 互动视频定义

上世纪末开始，随着数字电视逐步普及，人们不满足于被动收看电视，有了越来越多的互动需求。为了满足人们的互动需求，电视行业科技人员根据当时技术条件和发展趋势研发了若干“互动电视”技术标准，主要如下：

- 本世纪初，欧洲的 ETSI 推出了 MHP^[1]，MHP 主要基于单向广播的互动应用，考虑当时机顶盒 SoC 的处理能力，MHP 的应用分为 MHP-J（在 J2ME 上扩展电视相关 API）和 MHP-HTML（在 XHTML1.0、CSS2.0、ECMAScript 第一版基础上扩展电视相关 API）。随着技术的不断发展，MHP 目前已逐渐退出历史舞台，意大利等欧洲国家已经从原来机顶盒中间件的实现方式过渡到混合广播宽带系统。

- 2009 年欧洲推出 HbbTV^[2]（2010 年成为 ETSI 技术规范），目前在欧洲广泛应用。HbbTV 主要支撑基于 Web 的应用，并在 HTML5 规范上增加了电视相关的 API^[3]。HbbTV 的应用和数据可通过广播信道传输，也可通过互联网传输。

- 2013 年日本 NHK 推出 Hybridcast 服务，目前 Hybridcast^[4]已在日本广泛应用。Hybridcast 基于 HTML5 页面应用设计，可通过广播和互联网两条通道将节目内容和应用推送至终端的电视大屏和移动小屏上，Hybridcast 主要用于提供各类信息，包括新闻、股市、电子节目指南、视频点播以及与在线答题相关的节目信息。

- 2018 年 ATSC 正式发布 3.0 标准，在 ATSC 3.0 标准的第 A/344 部分，对互动内容（Interactive Content）进行了定义^[5]，目前 ATSC 3.0 中互动内容的主要应用方向有个性化节目单、基于用户观看行为的广告精准投放等。

- 2019 年 ITU 发布的技术报告 BT. 2267-10 综述了目前 5 种用户覆盖较多的、基于 HTML5 的综合广播宽带系统（包括 HbbTV、Hybridcast 等），并讨论了如何制作可运行于不同 IBB 系统的应用^[6]。

以上几种“互动电视”的主要特点是围绕“线性视频”开展内容互动业务。“线性视频”是指在用户角度，按照时间表播放的一个或者多个电视节目，用户只能收看按时间表播放的当前电视节目内容，无法进行暂停、回看、快进、快退等操作，例如：通过固定带宽信道（有线、卫星、地面）广播的电视节目，通过

IP 网络多播的电视节目等。互动信息通过带外（如互联网宽带）或者带内（打包在 TS 流内）传输，由于线性视频的特性，很难在“线性视频”上开展分支形式的互动内容。而“非线性视频”则通过双向网络单播，且可从用户角度进行暂停、回看、快进、快退等操作。

5G 高新视频-互动视频（特指通过 5G 等宽带网络向公众提供的互动视频节目服务，以下简称：互动视频）是指以“非线性视频”内容为主线，在“非线性视频”内容上开展的可支持时间域互动、空间域互动、事件型互动的**内容互动**视频业务，该业务具有分支剧情选择、视角切换、画面互动等交互能力，能够为用户带来强参与感、强沉浸度的互动观看体验。

互动视频的互动类型主要包括以下三种：（一）**时间域互动**，用户在观看过程中通过主动选择剧情分支可改变剧情发展，从而在时间轴上创造不同分支的内容走向，例如剧情选择。（二）**空间域互动**，用户在观看过程中可根据个人喜好，选择观看同一时间不同空间场景的内容，从而获得更加丰富的内容信息，例如视角切换。（三）**事件型互动**，用户通过特定行为（例如点击、滑动、长按等）触发视频中的人物、场景、物品等属性信息或关联的视频内容，从而扩展内容的深度，例如画面互动。

互动视频系统包括互动视频制作平台、互动视频服务平台、互动视频播放系统等模块，如图 1-1。

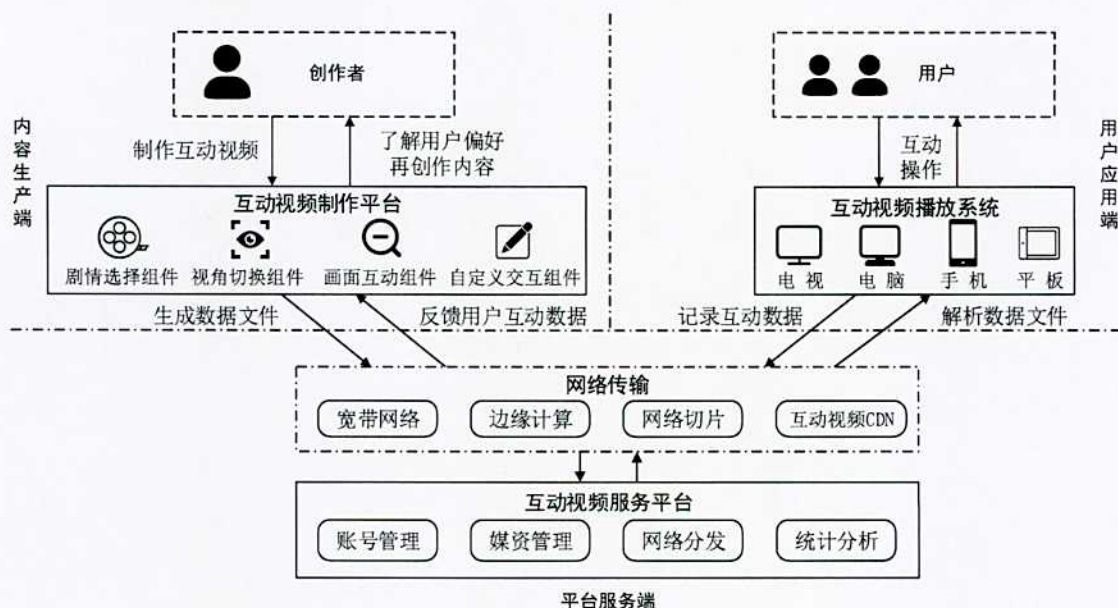


图 1-1 互动视频系统组成

创作者在互动视频制作平台完成互动视频制作并传输至互动视频服务平台，服务平台进行媒资管理，并将互动视频内容分发至互动视频播放系统，用户在播放系统上观看体验视频内容，同时播放系统收集用户互动数据并反馈至服务平台，供平台统计分析或指导创作者再创作。

互动视频制作平台通过互动视频制作工具为视频添加互动视频组件样式和参数，可编辑、预览视频互动效果，使视频具备互动能力。

互动视频服务平台通过互联网调用互动视频服务，完成互动视频的账号管理、媒资管理（转码、编目、审核、发布）、网络分发、统计分析等业务。

互动视频播放系统通过播放器和互动引擎实现互动视频的播放、缓冲、渲染、交互等功能，同时采集用户互动数据，供平台统计分析或指导创作者再创作。

5G 等宽带网络为互动视频制作提供大带宽、低时延的传输保障，边缘计算为互动视频处理提供算力支撑，网络切片为视频直播提供专有通道保障，互动视频 CDN 为视频内容分发提供交互体验支撑。

互动视频的数据结构包括播放区间和互动节点。一个完整的互动视频中包含多个视频片段，其中每一个视频片段所需相关信息的集合定义为一个播放区间。每一个交互内容信息的集合定义为一个互动节点，互动节点是互动视频内容播放控制的枢纽，一个互动节点中包含若干互动组件，每一个互动组件中定义了互动样式和参数，互动参数中配置有互动触发后的播放区间链接等信息。

1.2 互动视频发展历程

互动视频的核心特点在于交互式体验。从最初互动小说到互动影像游戏，再到 AVG 文字冒险类游戏，交互的方式在不同题材类型中迸发出不同的表现力。

1966 年，世界上第一部互动电影《lověk a jehodm》（英文译名 Kinoautomat）于捷克斯洛伐克拍摄完成，全片用 63 分钟的时间讲述了一名中年男子在其妻子生日当天，由于误解而导致的一连串离奇遭遇，每当剧情进入关键转折处，会由主持人给出红、绿两个按钮，让观众根据个人意愿做出选择，而剧情走向也会根据大部分人的选择结果继续发展。该片于 1967 年在加拿大世博会展出，其新颖的交互方式在当时引起了轰动。

1970 年，互动小说《选择你的冒险历程》系列丛书采用分支叙述的方式，让读者选择不同的选项并开始“冒险”，这种将选择权交给用户的形式，强调代

入感并引入剧情与用户间的实时互动，让彼时的游戏行业看到了市场。于是在上世纪八九十年代，真人互动影像游戏开始兴起。

1983年，Cinematronics公司发行了历史上第一部互动影像游戏《龙穴历险记》(Dragon's Lair)，该游戏引入了快速选项模式，使游戏玩家作为主角参与其中，在遇到机关时需在规定时间内快速做出行动选择。由于游戏玩法新颖，在当时获得了一批游戏爱好者的追捧。《龙穴历险记》开创了交互影像游戏的先河，而后游戏市场利用交互的形式，相继推出了各类互动游戏作品，例如《太空王牌》、《午夜陷阱》等。

进入21世纪后，互动影像游戏逐步衍生发展出科幻题材性质的叙事作品，例如《428：被封锁的涩谷》、《Her Story》、《地堡》、《夜班》等，用户在参与过程中不断搜集故事线索信息，并依据个人选择走向不同的结局。在Quantic Dream工作室推出的《底特律：成为人类》互动游戏中，玩家可以操控三位生化人来面对不同的剧情发展选项，选项的数值及选项组合将触发不同的事件走向，以极具刺激感的分支选项形式，迅速吸引了大量玩家。

1.3 互动视频国内外应用发展现状

1.3.1 国际互动视频应用发展现状

在互动视频的应用发展方面，Netflix、HBO、YouTube等视听网站均已布局互动视频业务，陆续推出了多部深受好评的互动剧，并在不断探索更多的互动形式和更佳的观看体验。

自2017年起，北美最大的影视视频供应平台Netflix开始尝试在一些动画中加入互动选择的形式，相关互动作品有《Puss in Book: Trapped in an Epic Tale》、《Buddy Thunder Struck: The Maybe Pile》等。

2018年1月，HBO上线了其首部互动剧《马赛克》，除电视、流媒体等常规播放模式，该剧还开发了同名App，用户可在App中进行交互式观影。

2019年10月，YouTube在线上了其首部互动作品《Markiplier大劫案》，该作品共包含61个视频片段和31种结局。同时，EKO、Ctrl-Movie等互动内容制作公司也在不断推出全新的互动视频内容。

在互动视频的技术标准方面，国际主流视听网站采用的互动视频技术方案均

不统一，实现架构上也存在较大差异。此外，截至 2020 年 7 月，ITU、ABU、ATSC 等国际标准化组织均未发布互动视频相关的标准规范、技术方案等。

1.3.2 国内互动视频应用发展现状

2019 年伊始，国内主流视频网站大力发展互动视频业务，推动其产业化发展。截至 2020 年 7 月，国内市场已推出百余部互动视频作品，互动视频制作机构、平台机构、应用机构呈现出百花齐放的发展态势。

2019 年 1 月，腾讯视频上线互动影视剧《古董局中局之佛头起源》，该剧设置了 4 个分支剧情、3 种结局。同年 9 月，互动影视剧《因迈思乐园》以每集 7-10 分钟的短剧形式出现，总时长约 60 分钟，并设置了 30 余个互动节点、3 种结局，在参与互动的用户中，人均互动 6.5 次，人均观看 2.8 个结局。2020 年 3 月，互动综艺节目《我+》在腾讯视频上线，节目中设置了 50 个互动节点、8 种结局，所有剧情总时长约 200 分钟，用户主动互动率达 30%。截至 2020 年 3 月，腾讯视频共上映了互动影视剧、互动综艺、互动电影等不同题材类型的 28 部互动视频作品。

2019 年 6 月，爱奇艺上线互动影视剧《他的微笑》，作为其首部落地的互动影视作品，该剧充分运用了剧情选择互动能力，设置了 21 个互动节点、17 种结局，所有剧情总时长约 200 分钟，最短的故事线仅 5-10 分钟，在参与互动的用户中，人均互动 9.5 次，人均观看 3.5 个结局。此外，综艺节目《中国新说唱 2019》中引入了剧情选择互动能力，并结合节目内容生产了互动视频广告，其用户主动互动率达 20%，广告总曝光次数提升约 18%。2020 年 1 月，爱奇艺上线影视剧《爱情公寓 5》中的互动剧集（13、35 集），如图 1-2（图片来源：爱奇艺）。其中第 13 集《弹幕空间》上线首日占据微博热搜榜第三，知乎热搜榜第二，截至 2020 年 2 月，总计观看用户量达 2821 万，其中 1514 万名用户观看了某一结局、492 万名用户观看了所有结局，总互动次数达 1.42 亿次，用户复看率达 81.9%（复看率指因次要结局产生重新观看的次数/次要结局出现的总次数）。



图 1-2 影视剧《爱情公寓 5》互动剧集

2019 年 1 月，芒果 TV 上线其首部互动影视剧《明星大侦探之头号嫌疑人》，该剧共 6 集，单集主线时长约 25 分钟，平均每集内有 20 余个剧情线索需用户互动参与寻找。

2020 年 1 月，优酷基于综艺节目衍生出其首部互动影视剧《当我醒来时》，该剧运用了互动视频中的剧情选择互动能力，共设置了 27 个互动节点、9 种结局，所有剧情总时长约 50 分钟，最短的故事线仅 6 分钟。

在互动视频的技术标准方面，2019 年 8 月，国家广播电视总局科技司正式立项制定《互联网互动视频数据格式规范》推荐性行业标准。该标准的主要目标是定义互联网互动视频的系统架构和数据格式，规范互联网互动视频的系统建设。2019 年 11 月，工业和信息化部科技司正式立项制定《互联网超高清视频播放软件 第 2 部分：互动视频技术要求》推荐性行业标准。该标准计划规定互动视频播放软件接口、功能构件等内容。

总体来讲，互动视频在国内市场仍处于起步探索阶段。在制作方面，相比于传统视频内容，互动视频因其多分支的故事结构，剧集制作的复杂度明显提高，创作者仍需探索和积累；在内容方面，目前市场上的互动视频作品整体评价较低，部分剧本内容仍不成熟，存在内容与互动脱节的问题；在技术方面，互动视频的内容制作流程、节目审核规范、播放系统要求、观看体验评价等方面缺乏统一的标准，大规模发展需要一定的技术和资金支撑。

1.4 互动视频应用场景

通过综合运用互动视频时间域互动、空间域互动和事件型互动中的多种互动能力，可为用户带来丰富、多元的视听体验。目前，国内市场已涌现出一批互动影视剧、互动综艺、互动短视频和互动影像游戏等视听内容，同时，互动视频已逐步应用于体育竞技、在线教育、电商购物和互联网广告等领域。

1.4.1 互动影视剧

借助互动视频时间域互动中的剧情选择能力，互动影视剧可设计多分支内容，用户在观看过程中可成为其中的“主角”并介入剧中环境，通过互动选择决定剧情的走向。通过用户互动与剧情内容的有机结合，提升观看过程中的视听体验。应用示例参看附录 A。

1.4.2 互动综艺

将互动视频时间域互动中的剧情选择能力引入综艺节目中，能够与不同题材类型的综艺节目结合迸发出新的创意，如竞猜问答类的综艺节目，在面对实时问题时，用户可从被动接收观看转变为主动参与作答。此外，通过引入互动视频空间域互动中的视角切换能力，能够使用户根据个人喜好自主选择观看视角，灵活“切换”机位。应用示例参看附录 B。

1.4.3 互动短视频

短视频具有短、平、快的特点，要求视频内容简短、易懂，因此短视频中的互动方式更简单、有趣。以科普类互动短视频为例，通过运用互动视频时间域互动中的剧情选择能力，设置不同选项的互动反馈，用户将更易理解科普知识、加深印象。应用示例参看附录 C。

1.4.4 互动影像游戏

互动影像游戏又称互动影游，是互动视频的一个重要应用场景。互动影游更侧重为用户提供游戏体验，借助互动视频时间域互动中的剧情选择能力，能够使用户深度参与到互动影游中，并影响其结局。应用示例如图 1-3（图片来源：互联网）。



图 1-3 互动影像游戏应用示例

1.4.5 文化旅游

对于文化旅游产业来说，互动视频为其带来了全新的线上“云旅游”模式，一方面能突破天气、场地等对传统旅游活动的限制，为游客带来全新体验和认识；另一方面能为旅游供应商提供更好的营销渠道和服务能力，从游前的行程预览、游中的体验服务到游后的行程回顾，借助互动视频的剧情选择、视角切换和画面互动等能力可全过程、全方位服务游览过程中的各环节。此外，在线下实景游览体验中，互动视频可紧密结合旅游景点承载的历史价值与人文情怀，推动中华历史文化元素全面融入旅游各个环节，助力文化和旅游融合，应用示例如图 1-4（图片来源：智令互动）。



图 1-4 文化旅游应用示例

1.4.6 体育竞技

在体育赛事等竞技节目中，经常存在多部摄像机同时拍摄的情况，但用户只能看到一个指定播出的画面。借助互动视频空间域互动中的视角切换能力，可将直播过程中多个机位的内容传送至客户端，用户可根据兴趣灵活切换机位，自主选择不同的观赛视角，使体育赛事的技巧性与观赏性完美结合，满足用户个性化、定制化的观赛需求，为用户提供更紧张、刺激的观赛体验，应用示例如图 1-5（图片来源：爱奇艺）。



图 1-5 体育竞技应用示例

1.4.7 在线教育

在线教育场景中引入互动视频，增加了教学方式的多元性，可以支持多场景教学环境切换和多终端教学模式互动。借助互动视频时间域互动中的剧情选择能力，学员可在线完成教员下发的考核任务，能够有效提高在线课堂的参与感，教员可根据学生互动反馈进行答疑。此外，教员还可对学员互动数据进行统计分析，了解学员的学习状况和诉求，便于后续教学方法的改进，应用示例如图 1-6（图片来源：智令互动）。



图 1-6 在线教育应用示例

1.4.8 电商购物

在互动视频中引入电商购物，可满足用户一站式的消费需求。用户在观看视频内容时，可能对节目中的某些商品感兴趣，在传统的消费路径中，用户需要先搜索掌握商品信息再进行采买，过程繁琐。借助互动视频事件型互动中的画面互动能力，用户可在视频播放过程中直接点击画面中感兴趣的物品，了解商品信息并完成订购，既简化了操作流程，又不影响观看体验，应用示例如图 1-7（图片来源：爱奇艺）。



图 1-7 电商购物应用示例

1.4.9 互联网广告

传统广告缺乏与受众的互动，并受播放时长限制。互动视频形态的互联网广告，借助时间域互动中的剧情选择能力，可有效激发用户对未解锁广告内容的好奇心，并打破传统广告的时长限制，用户可通过不同的剧情分支反复观看同一个广告不同方面的内容，提高广告内容的曝光率。此外，用户在互动中产生的行为偏好数据，可为广告方提供更精准的用户画像，便于广告的精准投放和营销。应用示例参看附录 A。

1.5 互动视频应用机遇

随着 5G 网络的覆盖普及，5G 网络所特有的“大带宽、低时延、广连接”特性为互动视频发展应用创造了条件，互动视频作为一种新兴的业务形态，通过互动参与可为用户带来浸润式的观看体验，有效解决了用户被动接收内容的弊端，为视听内容生产创造了新的可能性，给视听应用服务赋予了新的生命力。

在视听内容生产方面，用户在观看影视剧、综艺、短视频等视听内容过程中，存在多样性、个性化的观看需求，通过内容走向与互动方式的有机结合，可使内容更具深度和表现力。用户通过互动可获得更自主、浸润式的观看体验，产生更多的情感共鸣。

在视听应用服务方面，文化旅游、体育竞技、在线教育、电商购物等领域存

在各类互动需求，例如，在体育竞技领域，用户观看体育竞技节目的过程中，通过视角切换的互动能力，可主动选择个人感兴趣的镜头进行观看；在电商购物领域，将商品和服务从传统的图文宣传、视频宣传升级到互动视频宣传，用户通过互动可快速定位商品，商家通过用户互动反馈可感知其意向。

1.6 互动视频应用挑战

目前，互动视频在技术应用和业务发展过程中也面临一些问题，主要包括以下五个方面。

(1) 互动视频制作成本高。相比传统视频内容，因互动视频具有分支剧情，导致内容素材的拍摄量大幅度增加，内容制作成本也随之提升。此外，相比于传统视频的单线性叙事结构，互动视频的多分支叙事结构给内容制作全流程带来了全新的挑战，对内容制作人才也提出了更高的要求，进一步提高了制作成本。

(2) 受众群体覆盖率低。目前互动视频的受众群体偏小众化，观影人群中大部分用户仍处于“不愿互动”的状态。一些互动剧中预设的强交互要求会受到用户的抨击，如作品内容、互动设计无法吸引用户的眼球，容易出现高投入创作的互动视频，用户观看分支完成率却不高的情况。

(3) 互动视频可移植性差。在当前互动技术架构下，与视频内容相关的互动方式多有特定设计或仅能适用于某款播放系统，缺乏统一的数据格式、播放和接口规范与行业标准，目前难以实现跨平台复制和跨终端播放。

(4) 用户隐私保护有待加强。用户在参与互动过程中，互动视频播放系统将记录用户的互动行为并进行分析，在收集、整理、统计、分析用户信息的过程中，如未遵从相关用户隐私保护规定，可能导致用户隐私信息泄露或滥用，因此对用户隐私安全性的保护有待进一步加强。

(5) 网络承载能力有待提升。互动视频对网络抖动和网络时延均提出了更高的要求，尤其是在博物馆、运动场等人群密集场所应用时，将场馆内的互动视频内容传送给参与互动的现场观众，将对网络传输造成巨大压力，为确保回传网络的带宽和时延，可通过在承载网中部署 FlexE 切片等技术方案，提升网络承载能力，增加网络稳定性。

2. 互动视频关键技术

2.1 互动视频整体技术框架

互动视频整体技术框架由互动视频制作、互动视频编码和封装、互动视频传输和互动视频播放等部分组成，如图 2-1。



图 2-1 互动视频整体技术框架

在互动视频整体技术框架中，首先应进行互动视频的生产制作，然后进行互动视频信令数据和视音频流数据的封装压缩、智能编码等，并分发传输至用户端，最后用户端通过互动播放引擎观看体验不同互动类型的视频内容。

研究探讨互动视频在内容制作、分发传输、终端显示等环节的关键技术，对于提升互动视频的视听体验，具有重要的技术支撑作用，本章后续内容将重点探讨互动视频制播流程中的关键技术。

2.2 信令数据封装技术

互动视频封装的数据包括播放区间的视频流数据和互动节点的样式、配置参数等。互动播控信令（Interactive Playback Control Signal, IPCS）是互动视频的重要关键技术，是支撑互动视频的智能控制内核。IPCS 技术在互动视频制作平台的支撑下，对互动节点的界面、配置参数等数据进行封装。在互动视频播放器中，通过信令协议控制互动节点创建、配置互动节点样式、监听互动行为、

执行互动指令，并通过对互动节点的控制建立各视频片段之间的故事线，构建视频内容、互动事件、用户行为之间的智能关联。

(1) IPCS 的分层

IPCS 技术可抽象为三个层级：信令数据层 (Signal Data, SD)、信令展示层 (Signal User Interface, SUI) 和信令控制层 (Signal Control, SC)。

信令数据层规定目标互动的特征模型，确定目标互动的功能特征。通过构建完善的特征模型有助于准确构建目标互动样式与行为。

信令展示层完成信令数据的分析与提取，实现目标互动界面的绘制和展示。

信令控制层定义目标互动的行为能力，系统宜以 FIFO 模式完成控制指令的条件性响应。

IPCS 三层结构模型如图 2-2。

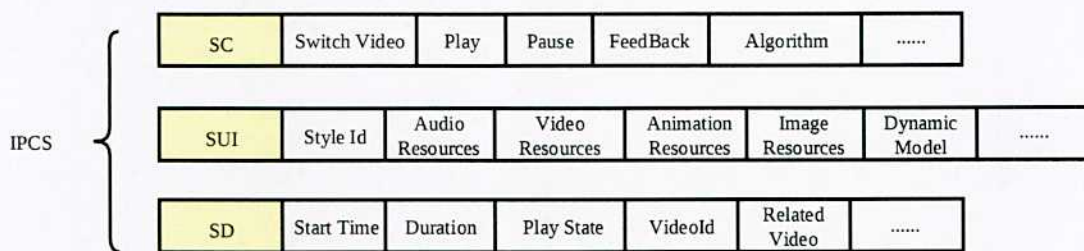


图 2-2 IPCS 三层结构模型

(2) IPCS 模型关键数据

IPCS 模型关键数据定义见表 2-1。

表 2-1 IPCS 模型关键数据

关键数据		说明
SD	Start Time	互动开始时间，以秒为单位（点播专用数据） 正值表示从视频片头开始计算 负值表示从视频片尾开始计算
	Event Id	互动事件标识（直播专用数据）
	Duration	互动持续时间
	Play State	互动开始时视频的播放状态：暂停或播放
	VideoId	互动节点所在的播放区间标识
	Related Video	表示与该互动节点相关联的所有视频片段
SUI	Template	互动组件样式模板 1: 剧情选择

关键数据		说明
		2: 视角切换
		3: 画面互动
		4: 自定义交互
	Audio Resources	互动节点所使用的音频资源
	Video Resources	互动节点所使用的视频资源
	Animation Resources	互动节点所使用的动画资源
	Image Resources	互动节点所使用的图像资源
SC	Switch Video	互动视频片段切换命令
	Play/Pause	互动视频播放控制命令
	FeedBack	互动反馈命令
	Algorithm	互动因子运算命令

(3) IPCS 生产及传输

通过互动视频制作平台向 IPCS 模型注入相关数据后，互动视频服务平台将对数据进行加密及高效压缩，然后完成 IPCS 数据与互动视频之间的双向数据绑定，并根据 IPCS 数据对视频进行二次生产，利用 IPCS 模型中的关键数据智能生产关键帧、视频片段过渡效果等。最后 IPCS 数据和视频内容经媒体资产管理系统完成编辑、审核后，上传至 CDN 进行存储和分发。

(4) 直播场景下 IPCS 技术的实现差异

IPCS 技术除可应用于点播场景外，还可在直播场景下使用，但在关键数据和互动启动方式上略有差异。由于直播场景下数据传输具有实时性，因此无法通过 Start Time 确定互动的开始时间，需使用 Event Id 作为互动的唯一标识。在直播场景下，经后台服务分析应启动互动时，将该互动事件的 Event Id 加入直播数据流中推送至终端，由终端处理展示。

2.3 多视点编码技术

在互动视频视角切换场景下，应重点关注以下两方面内容：

- 多视角拍摄与传输，对网络传输能力提出了更高的要求，因此需提升视频流端到端的编码和传输效率；
- 保证用户在视角切换操作后的观看体验平滑、流畅。

为降低互动视频业务的制作、传输、分发成本，提升用户互动体验，在视角

切换场景下，宜使用如下多视点编码关键技术。

(1) 时间同步技术

为保证多视点画面间的时间同步，可在前端编码的每一个分片中插入时间戳，在终端播放解码时，解析对应的时间戳，并通过视频预取、缓存的方式进行播放控制，以屏蔽因网络传输时延导致的画面不同步问题，业务流程示意如图 2-3。

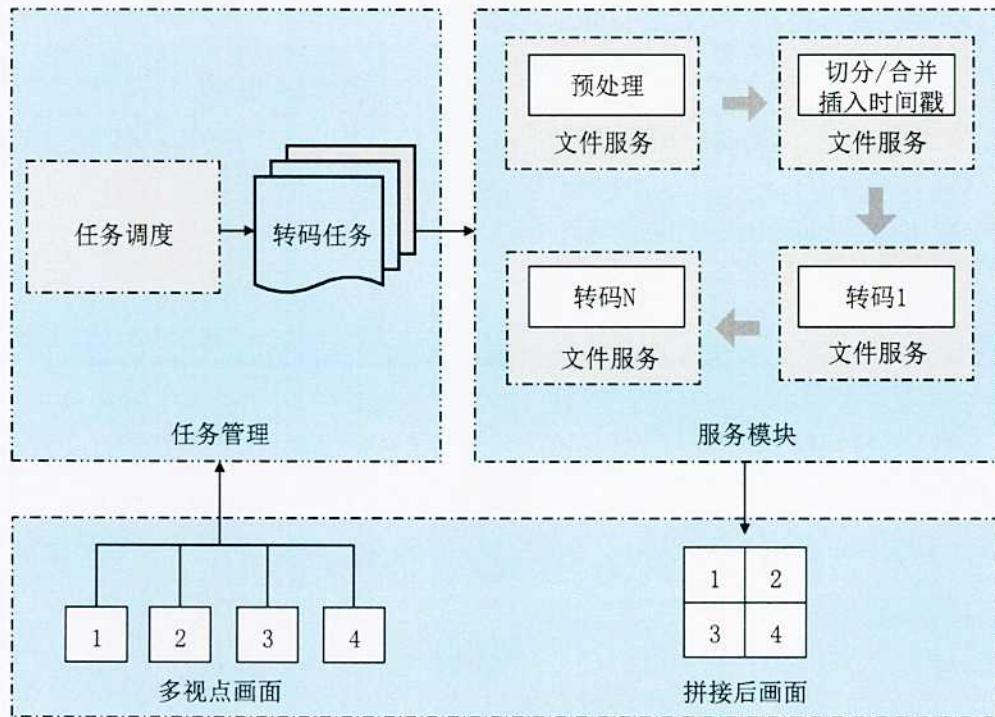


图 2-3 时间同步技术示意

(2) 编码效率优化技术

与单视点视频相比，多视点视频的数据量随摄像机数量的增加而线性增加，为解决数据存储量受限问题，ITU-T 和 MPEG 的联合视频组提出了多视点视频编码（Multiview Video Coding, MVC）标准，以充分利用视角间的冗余减少传输或存储的数据量。MVC 是未来视频通信领域中的一项关键技术，也是国际视频标准化组织正在研究的热点问题。

MVC 研究的关键技术包括：MVC 的预测结构、多视点视频的运动和视差补偿、多视点视频编码中的视点图像插值和颜色校正等技术。目前多视点视频编码主要基于 H. 265/AVC 编码框架而设计，如图 2-4，输入端由摄像机阵列拍摄的 N 路视点视频并行输入到 MVC 编码器，解码端根据用户的选择恢复出一个或多个视点的视频。相比 H. 265/AVC 标准，MVC 标准支持灵活的视间预测结构以挖掘视间相关性，因此编码效率显著提高。

为进一步提升面向终端传输数据的有效利用率,提升用户互动时的响应效率,在前端编码侧采用小切片技术,每一个切片可低至 200ms,确保所需分片可快速下载至用户终端;同时在编码时采用随机插入帧技术,保障不同视点的视频流可在 1~2 帧完成切换。

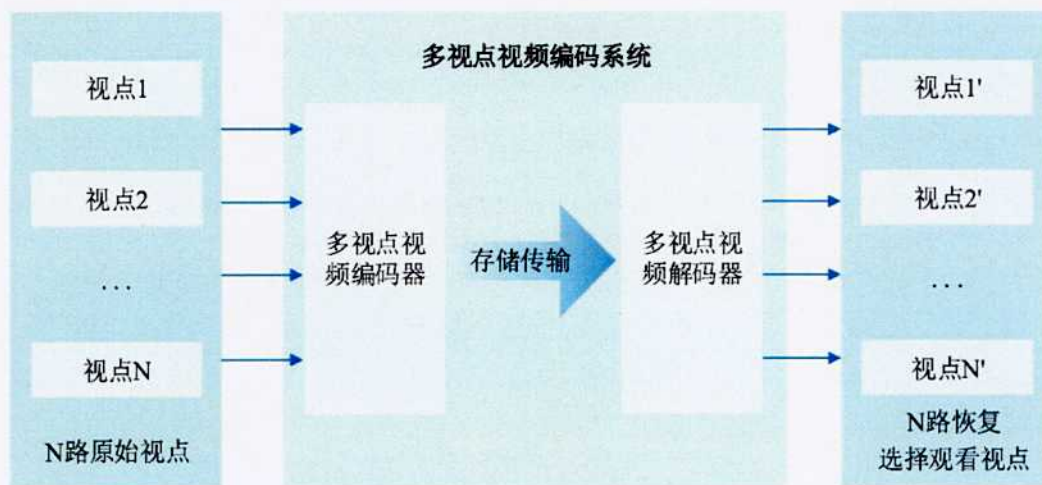


图 2-4 多视点编码技术框架

2.4 智能编码技术

当前,用户对视频清晰度、流畅度等观看体验的要求越来越高,各播出平台的 CDN 带宽成本也随之提升。智能编码技术旨在将编码比特放在最重要及人眼最关注的地方,在不降低、甚至增强视频源画质体验的基础上降低视频码率,减少用户 CDN 带宽成本,提升用户观看流畅度。

目前,已有较多结合深度学习、视觉识别技术优化视频压缩的方法,通过对视频画面类型分类、前背景分割以分配不同的压缩码率,进一步提升视频压缩效率。视频场景可分为游戏、秀场、体育、户外、动漫、美食、影视剧等十几个大类、几十个小类场景,通过视频场景分类实时识别,可根据视频源码率、帧率、分辨率、纹理和运动变化幅度等情况以及综合机器负载和画质效果,选择最优编码模板参数,不同的编码参数包括 IBP 帧类型、量化参数 QP、分辨率等,利用上述技术可实现互动视频的“提质降码”。

2.5 大带宽、低时延传输技术

在相同的视频格式和编码方式下,视频码率越大、清晰度越高,传输的数据

量越大。为保证互动视频播放过程中的清晰度和流畅性，传输带宽应符合一定的要求。例如 4K 显示屏对应的解析度为 3840×2160 分辨率，基带码率至少为 12Gbps，需要的传输码率（AVS2 和 H.265 标准压缩后）为 20~50Mbps。

在流畅性观看体验上，互动视频与普通视频的数据预加载要求一致，为达到不卡顿的播放体验，互动视频应进行数据预加载，而互动视频又存在一定的特殊性，视频内容可能由多视角画面组合播放，且一个视频片段后可能有多个待播放的分支视频，因此技术方案应考虑同时预加载多个视频片段。

不同用户的网络情况存在差异，在视频编码全过程中，平台需根据实际情况对多种清晰度（例如 480P、720P、1080P 等）进行编码。在 4G 网络情况下，用户可默认使用 480P 格式播放，以满足流畅播放的观影体验，在切换至 WLAN 时，可同步切换至 720P 或 1080P 等更高清晰度，在满足流畅观影的前提下获得更好的清晰度体验。因此，需要编码系统对视频片段进行多种清晰度格式编码，以满足用户根据实际网络情况在不同清晰度间平滑切换，具体切换过程如图 2-5。

在已实现多种清晰度的前提下，可采用自适应码流技术进一步优化播放效果。在传统的多种清晰度技术方案中，用户需自行根据网络情况调整清晰度，而利用自适应码流播放技术，系统能够智能判断网速变化，从而自动为用户选择最适宜播放的码流，实现当前网络条件下的最佳播放效果。随着 5G 技术的普及，5G 网络“大带宽、低延迟、广连接”的特性，将为用户提供更高码率、更加流畅的观影体验。

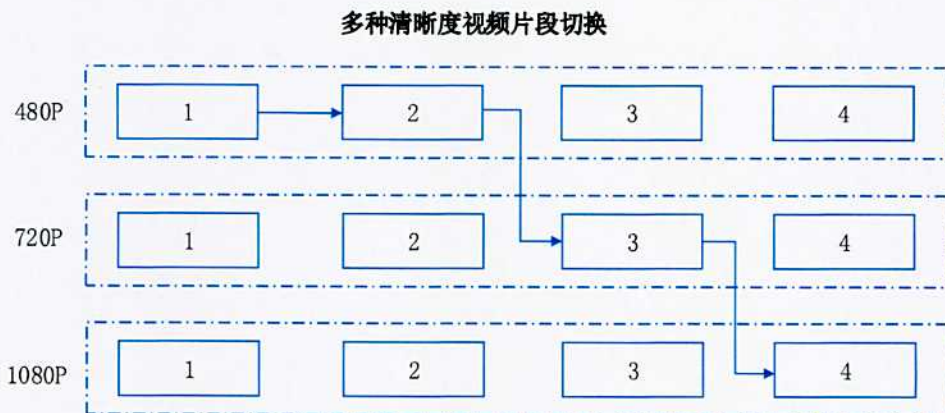


图 2-5 多种清晰度视频片段切换示例

2.6 网络切片技术

通过网络切片技术可以使网络运营商在同一硬件基础设施上切分出多个虚拟网络，实现网络资源按需分配、灵活组合的能力，以快速满足各种互动视频应用场景的业务需求，提升服务质量，优化用户体验。

此外，还可将 MEC 与 5G 网络相结合，提升互动视频在体育赛事直播场景下的时效性。MEC 是指在靠近人、物或数据源头的网络边缘侧，融合网络、计算、存储、应用核心能力的开放平台。利用 5G 网络可实现拍摄视频数据采集回传和反向控制，MEC 完成视频数据本地分流，同时完成视频内容制作，并向终端用户推流分发。

2.7 互动播放引擎技术

互动播放引擎是互动视频播放系统的核心部分，可为互动程序运行提供环境，为互动程序访问提供接口，并负责与播放器、终端设备及后台服务交互数据。互动播放引擎技术通过封装相关能力接口，可有效解决互动视频跨平台、跨终端播放兼容性的问题。互动播放引擎的技术架构如图 2-6。

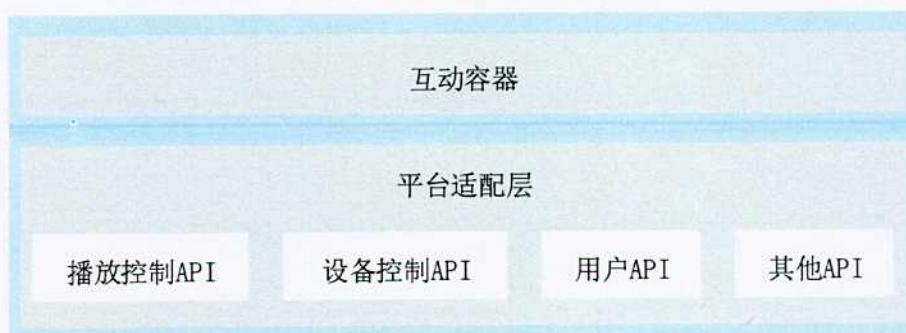


图 2-6 互动播放引擎技术架构

互动容器是互动播放引擎上层模块，可为互动组件提供运行环境。互动播放引擎下层模块包括播放器控制模块、终端设备控制模块、App 平台相关模块以及与用户相关的后台服务模块，这些模块通过网络调用、SDK、系统消息、环境变量等方式提供读写接口。

互动播放引擎将上述模块封装为四个 API 接口，其中播放控制 API 主要用于控制视频播放状态、播放进度、播放倍速，获取当前播放进度、播放状态、片段总时长等信息。设备控制 API 主要用于调节系统音量、设备亮度、应用触觉反馈

等。用户 API 主要用于账号管理、历史存档、数据采集等。其他 API 主要用于绘制故事线、监听互动操作及事件响应等。通过调用上述 API，互动相关程序可获得播放相关的控制能力、终端设备的交互能力和用户信息读写能力。

互动播放引擎的适配层可实现跨平台兼容，消除多平台底层模块调用的接口差异，为互动相关程序提供友好的运行环境。这种架构可较好的解耦互动控制逻辑和播放逻辑，达到灵活支持内容与用户实时互动的目的。

2.8 智能预测技术

通过人工智能技术的深度应用，互动行为的感知将更加精准和智能。用户观看互动视频时发生的互动操作、产生的互动结果等用户行为数据将传输至互动视频服务平台，服务平台通过对用户观看时长、观看频率、互动选择、浏览轨迹等行为数据进行统计分析和行为预测，可从用户模糊的行为习惯中精准地预测出用户喜好，以提升用户体验。

互动视频播放时，基于深度学习的推荐系统可将视频理解与用户画像相结合，并根据用户的观看行为、类别标签、地域属性等信息，建立多种类型的推荐模型。推荐模型大多包含输入层、模型层和输出层，输入层的数据主要包括用户的显式（或隐式）反馈数据、用户画像、互动行为、视频数据等；模型层通过机器学习技术进行数据分析，常用的模型包括卷积神经网络、循环神经网络等；最后由输出层输出内容推荐的预测结果。

2.9 5G 环境对互动视频关键技术的影响

（1）视频预加载不再是视频平滑切换的必要条件

使用时间域互动中的剧情选择能力时，应保证用户完成互动操作后，流畅、平滑地切换到后续视频片段，不影响观影体验的流畅性。但由于互动视频播放时，释放当前视频片段、加载下一视频片段的过程会产生一定的时间消耗，而该过程的耗时将影响播放的流畅性。为保证后续视频片段的快速起播，目前 4G 网络环境下，采用视频预加载技术对后续视频片段进行预加载提前缓冲，如图 2-7。

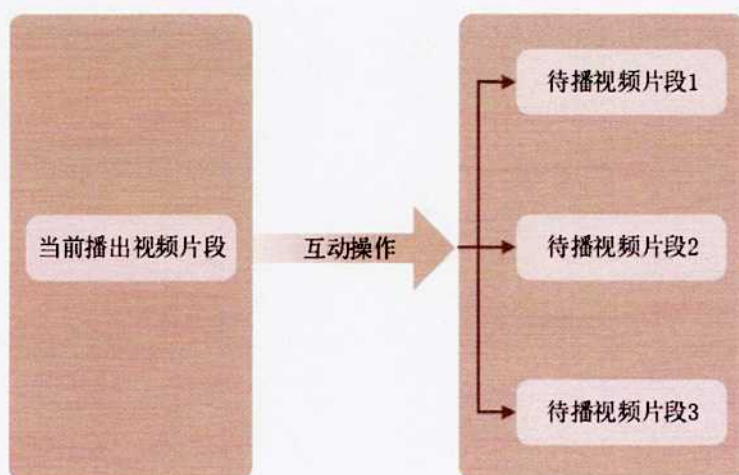


图 2-7 互动视频片段预加载示意图

随着 5G 网络的覆盖普及，5G 网络所特有的“大带宽、低时延、广连接”特性，使视频预加载技术不再是视频平滑切换的必要条件，用户做出任一选择，后续视频片段都可经过 5G 网络快速流畅起播。

(2) 多路视频流同时传输支持多视角灵活组合

使用空间域互动中的视角切换能力时，对于同一空间的内容，多个摄像机有多个拍摄视角，用户可根据个人喜好随时切换视角观看，要求对同一故事情节的内容，可平滑切换不同视角观看。在网速受限的条件下，当前技术方案如图 2-8。

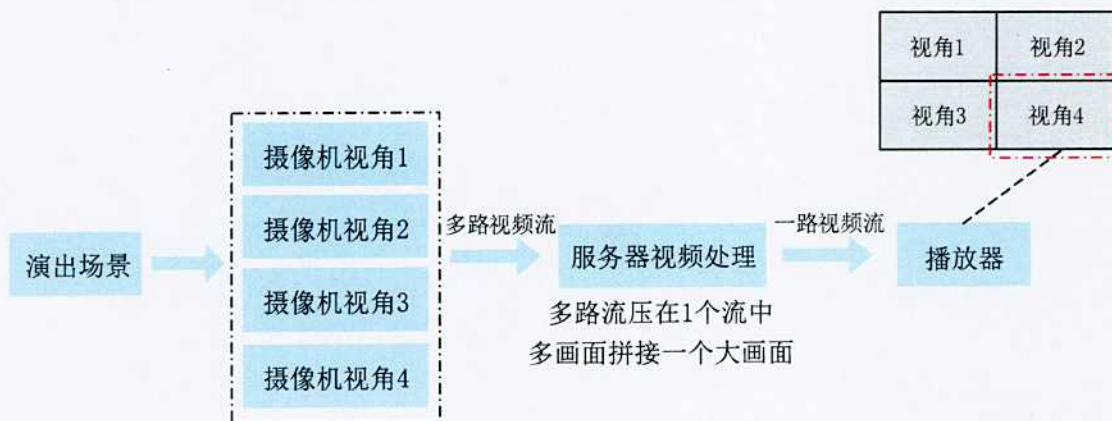


图 2-8 现有条件技术方案示意图

视频采用多个摄像机从不同角度拍摄，生成多个视频流（本例以 4 路视频流为例），视频处理服务器将 4 路视频流拼接为 1 路多画面的视频流，传输至播放器，播放器将播放框移动至不同的“子画面”上以实现视角切换的能力。

但上述方式不够灵活，需在视频处理初期进行压流，且切换的数量无法灵活

定制。随着 5G 时代的到来，多路压流将不再是唯一的处理方案，播放器可以同时接收多路视频流，实现真实切换多视角视频流。具体方案如图 2-9。

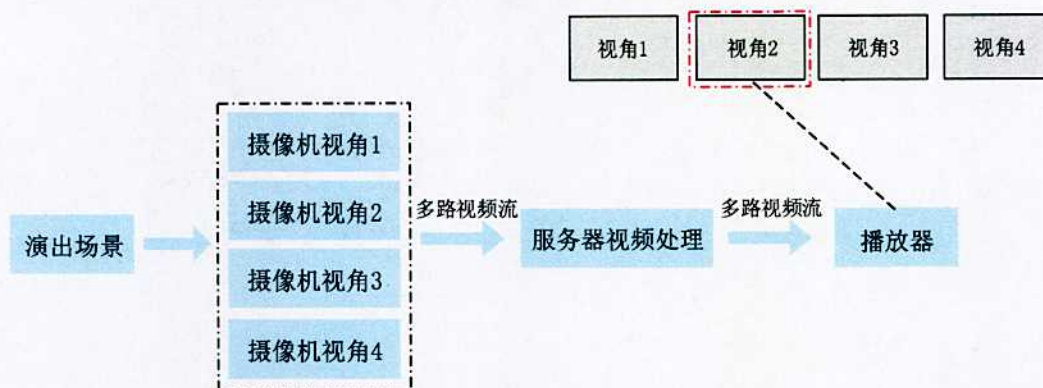


图 2-9 5G 环境下技术方案示意图

该方案可只传输其中一个视角的数据流，当选择另一个视角时，实时切换至另一个视频流，可能会导致毫秒级的延迟，但视角切换方案不再受限于前述方案。

3. 互动视频端到端解决方案

3.1 互动视频端到端系统架构

互动视频端到端系统架构，由互动视频制作平台、互动视频服务平台、互动视频播放系统组成，连接创作者端和用户端，如图 3-1。



图 3-1 互动视频端到端系统架构

PGC、UGC、MCN 机构等各类创作者，在互动视频制作平台为视频添加互动视频组件样式和参数，编辑、预览视频互动效果，使视频具备互动能力。

互动视频服务平台完成互动视频的账号管理、媒资管理（上传、转码、编目、审核、发布）、网络分发、统计分析等业务。并将视频内容分发至互动视频播放系统，通过对播放系统反馈的用户互动数据进行统计分析，优化剧情二次创作。

互动视频播放系统实现互动视频的播放、缓冲、渲染、交互等功能，并借助剧情选择、视角切换、画面互动等互动能力采集用户互动数据。

用户在不同终端观看体验互动视频内容，并通过遥控器点选、鼠标点击、滑动屏幕、语音命令等方式参与互动。

3.2 互动视频生产发布流程

互动视频与传统视频相比，在剧本创作、视频拍摄、剪辑处理、审核发布等环节均存在一定的差异。规范化的生产发布流程，将有助于创作者更高效便捷地制作互动视频内容。互动视频的生产发布流程包括互动视频编辑、预览、提交、

审核、发布等环节，具体如下：

（1）编辑：包括但不限于音视频上传添加、剧情配置、互动节点样式设置、互动命令设置等步骤；

（2）预览：根据创作者的编辑配置，模拟终端播放效果，实现播放区间预览及完整互动剧情的交互式预览；

（3）提交：对编辑后的视频进行规范性校验，校验通过后，进行封装、加密等处理，并提交至互动视频服务平台进行审核；

（4）审核：对已提交的互动视频进行内容质量及安全审核，可采用机器审核、人工审核等方式；

（5）发布：将审核通过的互动视频在互联网播出平台发布。

3.3 互动视频创作流程

互动视频的创作流程在传统视频创作的基础上有所继承和发展，下面以互动影视剧的创作流程为例，介绍互动视频创作的核心环节及内容。

3.3.1 内容创作

互动视频的内容创作应在遵循传统影视创作逻辑、经典戏剧创作规律的基础上进行更具体系、更严密的逻辑设计，首先，应确定视频内容的逻辑框架，明确互动节点位置及相应的互动方式，同时考虑互动前后内容自然、流畅衔接的问题。

（1）确定视频逻辑框架

根据内容的关键事件，应确定互动节点位置及相应的互动方式，输出视频内容的故事线结构图。在设定互动节点位置时，应考虑互动事件对内容节奏的影响，保证互动与内容节奏间的平衡。

（2）确保内容衔接合理

互动视频播放至互动节点时，将向用户开放互动能力（例如：用户通过剧情选择可影响内容走向），应确保互动操作前后内容衔接自然、合理，使用户轻松、自然地完成互动，并快速理解后续剧情。

3.3.2 现场拍摄

现场拍摄阶段，除传统的视频拍摄注意事项外，还应对互动操作前后的画面内容、转场镜头进行特定设计，确保内容衔接的自然性与合理性。

3.3.3 视频制作

视频制作阶段，应在互动视频制作平台为视频添加互动能力，并完成剪辑、调色、音效、特效处理等工作，输出互动视频内容。

(1) 互动组件易感知

互动组件应遵循“一个界面，一个任务”的原则，互动组件应契合前后衔接的视频内容，通过简洁、易懂的文字表述，引导用户完成互动操作，以观看不同剧情、不同视角的内容，应避免过于主观的刻意引导。

(2) 互动界面合理

互动界面包括互动组件界面、播放器界面等。互动界面原则上应遵循简洁、合理的设计原则，一方面，需契合视频内容与画面风格，另一方面，在不同的素材画面下，需考虑界面整体的视觉效果与合理性，避免用户对视频内容造成误判。

互动综艺、文化旅游、体育竞技等应用场景的互动视频创作流程与互动影视剧的创作流程大体类似，同时，也应结合各应用场景的具体特点进行针对性的创作。

针对互动综艺内容，在内容创作阶段，应根据综艺节目本身的特性，定制化设计与综艺内容相符的互动形式；在视频制作阶段，应通过后期剪辑、视觉包装等形式将节目内容和互动能力进行有机融合、无缝对接，提升综艺节目的互动参与感和趣味感。

针对互动视频在文化旅游领域的应用，在内容创作阶段，应重点考虑利用画面互动、剧情选择能力增强用户的身临其境感，可在视频内容中增加文化知识类问答内容，并通过画面互动的形式引导用户进行作答，在提升用户参与感的同时，弘扬传统文化。

针对互动视频在体育竞技领域的应用，在内容创作阶段，应针对不同体育赛事的观赏特性提前规划引入相应的互动能力，例如：在足球比赛中，可利用剧情选择、画面互动能力，引入比赛结果预判和进球红包环节，提高球迷的观赛热情；在高山滑雪比赛中，可利用视角切换能力，实现运动员视角、周边视角、自由视角及全景视角等多视角观赛，满足用户个性化、定制化的观赛需求。此外，在视频制作阶段，应注意对互动相关的内容进行解说配音、字幕包装等，确保比赛内容和互动能力有序衔接。

3.4 互动视频制作平台

互动视频制作平台面向创作者，主要为创作者提供素材管理、故事线编辑、互动组件管理、互动应用模板、互动视频预览、互动视频提交等功能，便于创作者创作和管理互动视频。互动视频的创作流程包括如下环节：

- (1) 新建互动项目：创建一个互动视频工程或打开一个已有工程；
- (2) 上传视频素材：将视频片段和文字、图片、音视频等素材上传至互动视频制作平台；
- (3) 创建互动剧情结构：在故事线画布区域设计互动视频的故事线逻辑结构；
- (4) 编辑互动信息：为各互动节点选择互动模板，配置互动参数；
- (5) 预览确认：预览创作完成的互动视频，模拟互动，确认效果；
- (6) 提交项目：将互动视频项目提交至互动视频服务平台，进行审核、发布。

3.4.1 素材管理

素材管理是互动视频制作平台的基本功能，主要功能是协助创作者上传和管理图片、视频、音频等素材内容，主要分为上传和管理两部分。

(1) 上传

- a. 创作者制作并上传素材内容，系统支持 MP4、AVI、MXF、PNG、JPG 等常见的素材格式；
- b. 视频由互动视频制作平台传送至服务平台，经审核、转码等处理操作后，生成相应的 VideoId，回传至互动视频制作平台；
- c. 互动视频制作平台收到服务平台返回的 VideoId，保存项目关联的 VideoId，用于后续互动视频编辑、更新；
- d. 互动视频服务平台完成转码、审核等操作后，会同步至互动视频制作平台，创作者在互动视频制作平台可查看视频当前流程状态。

(2) 管理

- a. 创作者对已上传的视频，可进行基本编辑、查询、删除等操作；
- b. 创作者可查看当前视频的状态。

3.4.2 故事线编辑

故事线主要用于描述互动视频的内容结构和脉络。

(1) 故事线画布区

这是创作者设计互动视频剧情最主要的操作区，主要支持以下功能：

- a. 支持将素材管理中的视频拖拽至画布区，生成对应的视频播放区间；
- b. 支持以添加互动节点的方式将相关视频播放区间连接生成剧情分支；
- c. 支持以表单或其他形式设置互动节点的互动组件样式和参数。

(2) 故事线管理

主要完成互动视频制作平台的故事线逻辑处理，主要支持以下功能：

- a. 数据逻辑封装模块，处理创作者在编辑互动视频过程中产生的业务逻辑，包括上文提到的视频节点、分支、组件等；
- b. 视频版本比较模块，创作者编辑互动视频过程中会产生多个版本或多次历史修改记录，可支持创作者对比查看互动视频的变更历史。

(3) 故事线数据存储

用于存储互动视频的素材、章节、故事线及互动组件等相关数据，宜分为两个存储数据库。

- a. 草稿数据库。用于管理和存储创作者配置的所有互动数据，草稿数据库的数据不宜实时对外生效；
- b. 正式数据库。创作者正式发布互动视频后，互动视频数据会从草稿数据库同步至正式数据库，此时用户即可观看最新配置的互动视频。

为保障创作过程记录可追溯，需要建立过程数据缓存，用于记录创作者在制作过程中产生的互动数据、操作流水等历史记录。

3.4.3 互动组件管理

目前，常见的互动组件包括剧情选择组件、视角切换组件和画面互动组件等。创作者可在互动视频制作平台选择互动组件样式模板并配置参数，样式模板指互动能力的界面样式，参数指基于样式模板的标题、倒计时控件、按钮等信息。

互动组件可分为如下两类：

(1) 标准组件

互动视频制作平台宜提供剧情选择、视角切换和画面互动三种标准组件，创

作者可以直接调用并配置其参数信息。

(2) 自定义组件

自定义组件指区别于剧情选择、视角切换、画面互动三种组件的其他互动组件。该组件类型允许用户自定义配置互动组件的样式模板和参数。

3.4.4 互动触发条件

(1) 时间触发

常用于时间域互动类型中。例如：依据内容情节，设定视频中的互动开始时间，内容播放至互动节点时触发互动，同时视频画面提示用户进行互动，用户基于自身立场或偏好进行互动并决定内容走向。

(2) 行为触发

常用于空间域、事件型互动类型中。常见的行为触发方式包括：单次点击、连续点击、滑动、滑动特定路径、长按、摇一摇等。例如：用户在直播时可在多个视角间主动自由切换，进而在不同物理空间以不同视角来获得不同的观看体验。此外，用户在观看时可主动与视频内容发生交互，触发视频中的人物、场景、物品等属性信息或关联的视频内容，从而扩展内容深度。

(3) 数值触发

常用于时间域互动类型中。互动视频服务平台记录用户在较长时间跨度内一系列互动操作的数值，该数值可在一定程度上反映出用户的喜好和兴趣，通过不同数值可触发播放不同的视频内容或隐藏的互动事件。

数值触发包括两种类型，具体如下：

- a. 通过数值改变视频内容，使用户在特定时间内体验不同的视频内容、结局；
- b. 通过数值改变同一互动事件内的交互内容，使不同用户在同一互动事件中有不同的互动操作。

3.4.5 互动应用模板

互动视频制作平台宜为不同场景提供快速创作的模板，提高制作效率、降低制作门槛。例如：电商购物、知识问答等应用场景有着行业共性，互动视频制作平台宜提取出应用模板供创作者快速调用；同时，创作者也可自行制作模板供个人或创作者使用。通过调用应用模板，创作者可在几分钟内完成互动视频的创作，而无需接触复杂的视频编辑器，极大提高创作效率和质量。

3.4.6 互动视频预览

为使创作者便捷地创作互动视频，互动视频制作平台宜提供互动视频的实时编辑与预览功能，并可在两种模式下灵活切换。

(1) 编辑模式

创作者可对视频片段和互动节点进行编辑操作，对每个节点的样式、属性、操作等进行所见即所得的配置。

(2) 预览模式

创作者可对编辑完成的视频内容进行预览，并验证互动操作的可行性，便于创作者体验效果、发现问题。

3.4.7 互动视频提交

互动视频制作平台应提供互动视频提交功能，使创作者制作的互动视频能够实时提交至互动视频服务平台进行审核并发布，提交时宜支持配置发布标题、发布描述、视频分辨率等视频信息。

3.5 互动视频服务平台

互动视频服务平台通过互联网调用互动视频服务，完成互动视频的账号管理、媒资管理（转码、编目、审核、发布）、网络分发、统计分析等业务。互动视频服务平台宜提供各种开放服务的能力，通过直接调用这些能力以保障互动视频的观影逻辑，互动视频服务平台主要包含下述模块：

3.5.1 媒体资产管理

互动视频播放依赖于播放区间中的视频片段和互动节点中的组件样式、配置参数，及文字、图片、音视频等资源素材。媒体资产管理应制定规范化的处理流程，该流程主要包括视频上传、转码、审核、标准化和发布等环节，如图 3-2。

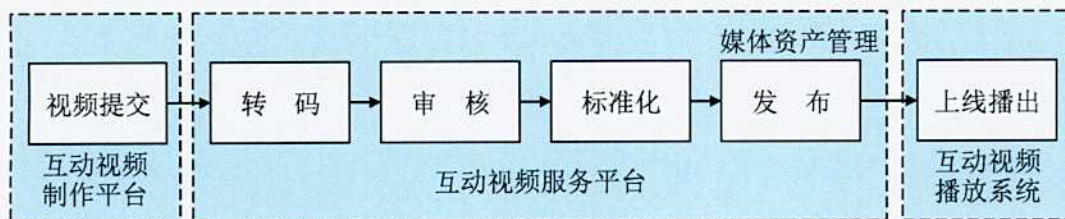


图 3-2 媒体资产管理流程

(1) 素材上传及转码处理

素材包括文字、图片、视频、音频、脚本文件等，通过互动视频制作平台上传至互动视频服务平台后，首先将图片、音视频等资源压缩、格式转换等操作处理，然后调度模块将转码请求发送至消息队列，转码完成后同步至 CDN 服务器。

(2) 内容审核及标准化

互动视频转码完成后，应对视频内容进行审核，审核流程如下：

a. 视频片段审核：需保证内容合理、合规，应从暴力、色情、反动、恐怖等多方面对视频内容进行审核，充分保证视频内容合法合规。如审核通过，则进行下一步骤；审核未通过，则通知创作者修改后重新提交；

b. 互动节点审核：应审核故事线结构及所有互动节点的配置信息。对于应用剧情选择互动能力的内容，其前后衔接应具有逻辑合理性，且所有分支路径最终均应走向内容结局，不出现某一支无限循环播放的情况。如审核通过，则进行下一步骤；审核未通过，则通知创作者重新修改内容结构及互动信息配置；

c. 视频片段、互动节点审核均通过，则自动触发内容发布上线。

互动视频审核可利用人工智能技术进行机器审核，而后再进行人工审核，如图 3-3（截图）。

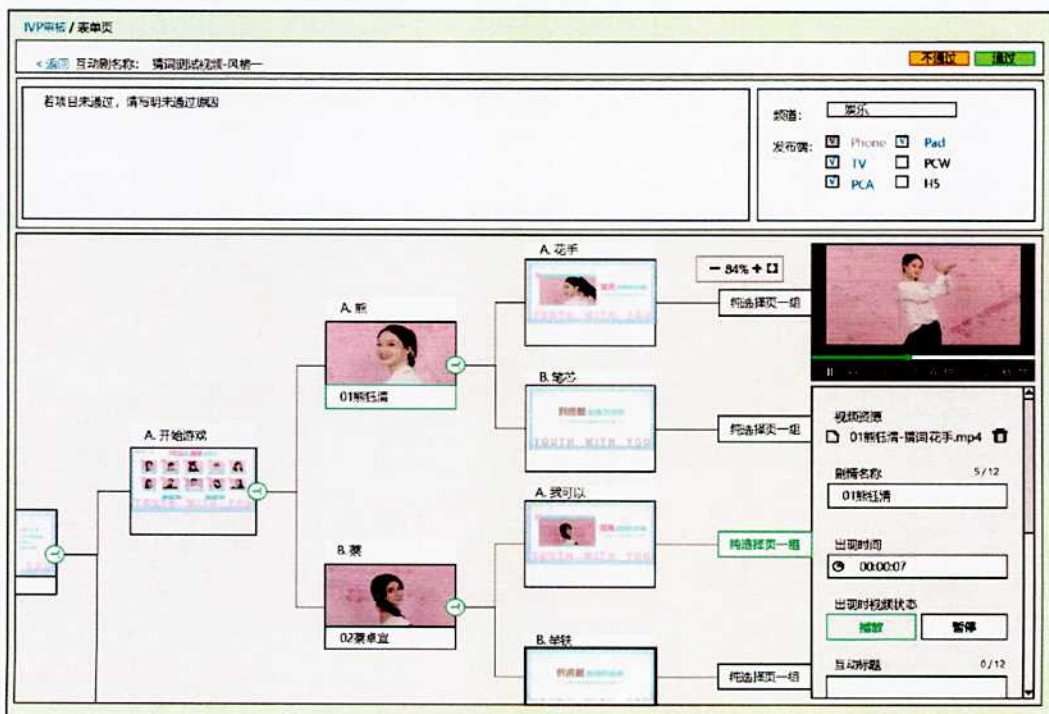


图 3-3 互动视频审核端示意图

通过审核的互动视频将经过标准化处理，完善基本信息、分类标签，建立分词索引等，为互动视频服务平台后续的推荐、查找等操作进行预处理。完成标准化处理的视频可发布至播放系统。

(3) 视频发布

互动视频服务平台对通过审核的互动视频在电视、电脑、手机、平板等播放终端进行排期发布。同时应制定合理的发布、下线规则，以保证业务稳定运行。

3.5.2 网络分发

互动视频服务平台可具备同时将视频片段、互动组件及其脚本文件上传至 CDN 中进行存储和分发的能力。

当前视频网络场景中普遍存在因带宽不足、时延过高等导致视频体验不佳的问题，归其原因在于区域热度内容未能分发至更接近用户的位置。CDN 内容分发网络将用户需要访问的内容分发到距离用户最近、网络服务质量最好的节点，同时通过负载均衡服务器自动地将数据调度到相应的高速缓存服务器，有效缓解互联网拥塞状况，提高用户访问的响应速度。

在提供边缘计算与存储能力时，可采用 AI 智能预分发推荐算法，通过监督/非监督学习、深度学习、时间序列等机器学习方法预测内容热度，并利用空闲时间段（凌晨）带宽同步更多的热点内容至边缘节点，错峰网络高峰值，减少回源带宽，有效降低骨干网负载。

互动视频从拍摄到网络分发的过程如图 3-4。

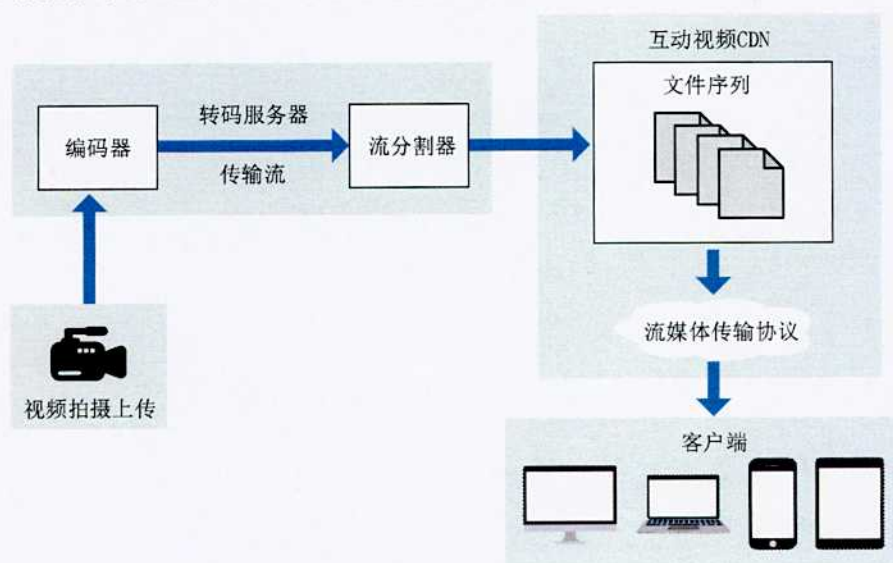


图 3-4 互动视频网络分发过程示意图

为保证互动视频多视角场景下的画面同步，及未来 8K 环境下用户的观看体验，互动视频 CDN 宜具备以下能力：

(1) 支持 4K 超高码率内容分发，为保证画质，应可支持超过 50Mbps 的视频码率；未来支持 8K 环境下，应可支持超过 100Mbps 的视频码率；

(2) 在分发 4K/8K 内容时，可通过 FOV 技术降低对终端解码能力及带宽的依赖，同时 4K/8K FOV 出流时延应小于 40ms，以满足端到端时延小于 200ms 的体验，避免用户产生“眩晕感”；

(3) 通过小分片处理技术，实现手势跟随滑动时延小于 200ms，实现“画随指动”的效果。

3.5.3 用户账号管理

(1) 用户账号

在视频播放应用中，系统通过用户登录的账号 ID 来实现某些个性化逻辑。而互动视频则具有一定的特殊性，每一部互动视频通过互动引擎程序来实现视频内容个性化逻辑，即互动视频间的个性化逻辑均不同。为使多个互动视频间互不影响，降低耦合性，宜按该思路进行账号体系设计，即同一个用户在不同互动视频的账号 ID 不同，针对每部互动视频单独派生一套账号。

(2) 创作者账号

创作者是上传视频、制作互动视频的用户，对于这部分用户 ID，应做好权限管理工作。系统宜给出细颗粒度权限定义，例如：上传素材、修改文案、删除视频、提交审核、仅预览等项，对应角色有开发者、运营者、超级管理员等。角色与视频发布也有一定的耦合关系，可在互动视频生产流程中定义审核发布权限。

3.5.4 历史观看存档

互动视频服务平台可对用户观看视频时产生的观看记录、互动结果进行管理，可实时查看观影记录、互动结果，并支持更新、删除等。当用户请求历史观看记录时应下发该用户的最新观看记录。

区别于传统视频，互动视频服务平台记录用户在某个互动视频中的观看状态，以及不同内容分支的观看进展尤为重要，以便互动视频起播后，互动程序直接定位至该用户的最近观看分支。历史观看存档的数据流程如图 3-5。

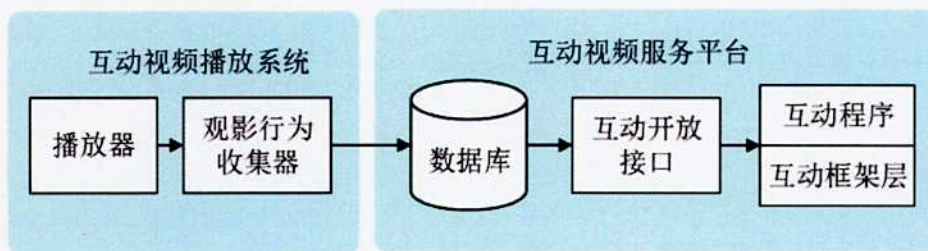


图 3-5 历史观看存档的数据流程

历史观看存档写流程：播放器向观影行为收集器上报播放信息，后台服务遍历索引信息：用户 ID→互动视频 ID→最近观看视频 ID。

历史观看存档读流程：互动程序通过 API 调用互动框架层，互动框架层调用互动开放接口，读取最近观看视频 ID 及播放时间信息并基于上述信息进行互动视频的播放。

3.5.5 数据采集分析

数据采集分析模块收集用户的互动反馈信息及视频播放统计信息等，有利于详细分析互动作品的目标受众，以及互动作品的优劣。数据采集功能的实现依赖互动视频播放过程中各关键点互动数据的上报，应在互动节点重置、互动组件展示、互动操作、视频播放结束等时机上报数据，见表 3-1。上报数据应包含播放区间标识、互动节点标识、播放时间等信息。

表 3-1 数据采集信息上报时机

上报时机	描述
互动节点重置	当用户在故事线中选择互动节点或重新播放视频时上报
互动组件展示	当互动组件开始展示时上报
互动操作	用户操作互动组件或互动组件倒计时结束自动选择时上报
视频播放结束	当互动视频全部播放完毕时上报

用户观看互动视频时发生的互动操作、产生的互动结果等用户行为数据将传输至互动视频服务平台，服务平台可提供数据传输接口，通过对用户观看时长、观看频率、互动选择、浏览轨迹等观看行为数据进行统计分析，了解用户行为偏好、兴趣喜好，以优化内容创作。数据分析主要对时间域互动和事件型互动两种类型的互动数据进行统计分析，宜按如下流程进行统计分析。

(1) 时间域互动（以剧情选择为例）

a. 互动视频数据分析维度统计，见表 3-2。

表 3-2 互动视频数据分析维度统计（剧情选择）

数据分类	指标
基础播放数据	总播放人数/总播放量/播放时长
	传统互动量（评论、点赞、弹幕）
	热度
互动人群分层数据	展示互动人数（该人群播放时长、播放次数）
	主动互动人数（该人群播放时长、播放次数）
	达成任一结局人数（该人群播放时长、播放次数）
	达成多结局人数（该人群播放时长、播放次数）
	达成所有结局人数（该人群播放时长、播放次数）

b. 解析互动脚本、获取相关数据记录；

c. 按分析维度计算各项数据指标，生成互动视频统计分析图。

(2) 事件型互动（以画面互动为例）

a. 互动视频数据分析维度统计，见表 3-3。

表 3-3 互动视频数据分析维度统计（画面互动）

数据分类	指标
基础播放数据	总播放人数/总播放量/播放时长
	传统互动量（评论、点赞、弹幕）
	热度
互动人群分层数据	展示探索点人数（该人群播放时长、播放次数）
	点击探索人数（该人群播放时长、播放次数）
	探索完整内容人数（该人群播放时长、播放次数）
	探索中途退出人数（该人群播放时长、播放次数）

b. 解析互动脚本、获取相关数据记录；

c. 按分析维度计算各项数据指标，生成互动视频统计分析图。

3.6 互动视频播放系统

互动视频播放系统运行在互动播放引擎之上，主要包括播放器、互动播放引擎和互动组件三部分，可实现互动视频的播放、缓冲、渲染、交互等功能，同时采集用户互动数据。互动视频播放系统技术架构如图 3-6。

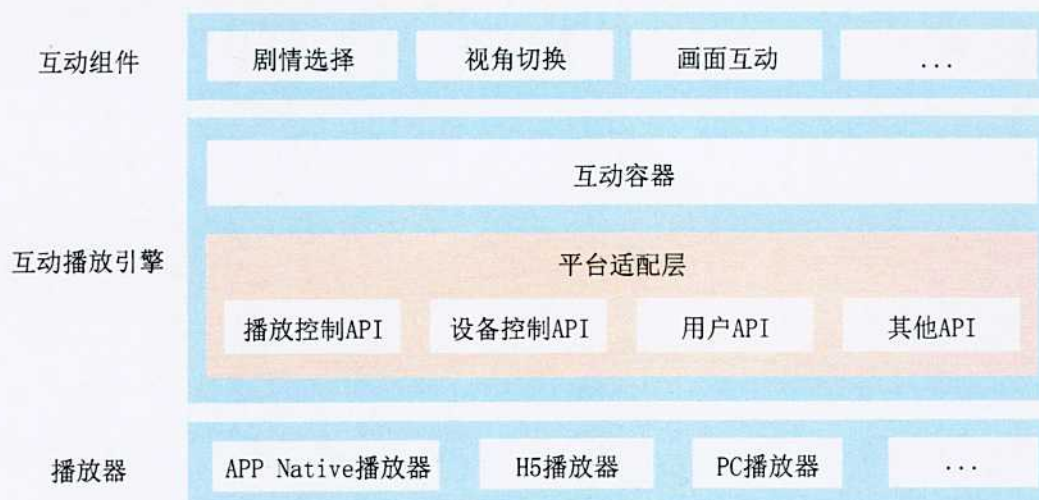


图 3-6 互动视频播放系统技术架构

3.6.1 播放器

播放器将视频文件解码并渲染出图像和音序列，并通过基本交互组件完成人机交互的功能，例如开始播放、暂停、拖拽、快进等。播放器是互动视频播放系统基本组成部分。

互动视频播放器可运行在多种平台场景，例如 APP Native 播放器、HTML5 播放器、电脑客户端播放器等。

3.6.2 互动播放引擎

互动播放引擎集成互动视频播放控制 API、设备控制 API、用户 API 及其他 API 等，为互动程序提供运行环境。互动播放引擎应具备表 3-4 列出的相关 API，并通过各功能模块调度上述 API 实现完整的互动过程。

表 3-4 互动引擎相关 API 列表

能力分类	主要能力项
播放控制 API	切换视频
	播放/暂停/播放进度

能力分类	主要能力项
	设置下一个播放视频
	获取播放时间/视频时长
	获取视频尺寸/信息
	视频片段预加载
设备控制 API	系统音量/亮度
	重力感应
	应用触觉反馈
	语音设备控制
	GPS 信息
用户 API	账号管理
	历史存档
	数据采集
其他 API	监听互动操作
	绘制故事线
	执行互动指令

互动播放引擎解决方案还应考虑以下几点问题：

(1) 平台适配层

主要解决跨终端和平台播放的差异性问题。同一互动视频既可在 iOS、安卓 Native 等手机终端播放，又可在 Web、电脑等平台客户端播放。不同终端和平台播放器的播放控制 API 不同，互动播放引擎的平台适配层应消除这些差异。

此外，不同终端和平台支持互动视频的能力不一，例如获取用户信息、设置/获取屏幕亮度、设置/获取屏幕旋转、设置/获取系统音量、历史观看记录、视频收藏等能力，宜通过封装统一的接口，保证互动视频在不同场景下体验一致。

(2) 互动后台封装

互动后台为互动视频内容和交互提供了功能接口，互动播放引擎需对互动后台的接口能力进行封装，并提供给互动组件使用。

(3) 互动容器

指互动组件的运行环境，在不同的终端和平台其运行环境不同。互动播放引

引擎应根据具体的终端或平台，提供不同的运行环境，并给运行环境提供播放控制 API、设备控制 API、用户 API 和其他 API 等，使互动组件可以无缝地运行在互动容器中。

3.6.3 互动方式

目前视频播放的终端设备包括电视、电脑、手机、穿戴设备等产品。由于各播放终端的交互输入方式不同，因此在不同播放终端上用户与互动视频间存在不同的互动方式。

电脑端主要输入方式为鼠标点击、键盘输入等，其播放的互动视频可通过鼠标点选视频画面完成互动操作。

手机端除手指点击外，还存在滑动屏幕、重力感应、地理定位、摄像头拍摄等操作输入方式，故可拓展更多的互动方式，例如：单次点击、连续点击、滑动、滑动特定路径、长按、摇一摇、手机姿态交互、语音命令等。

电视端主要输入方式为遥控器点选、语音交互等，适合以剧情选择、视角切换为主要互动能力的视频内容。此外，还可通过手机 APP 与智能电视互联，允许用户通过手机端的互动操作实现对电视端的内容遥控，使电视端也可支持点击、滑动、长按、摇一摇等更多互动方式，进而实现“多屏互动”。

穿戴设备相较于手机，与人体的动作状态联系更为紧密，互动方式更加丰富，例如：通过手指/手柄点击，人体姿态交互、语音命令等，适合更加沉浸式的互动视频。

3.6.4 用户操作响应时延约束

用户在进行互动操作时的反馈是否及时、流畅，是互动视频体验中重要的组成部分，因此，用户在每次互动操作后，系统实际反馈的时长需控制在合理范围内，以减少用户的等待时间，优化用户体验。用户在观看互动视频时的时延主要可分为等待类时延、播放类时延和互动类时延三类。

等待类时延，指用户从“接收信息→结合剧情分析判断→做出选择”的平均耗时，经统计该时延约 4s，建议 4s 作为互动节点设置的最短等待时长。

播放类时延，指视频起播、视频片段切换、用户拖拽进度条等条件下产生的时延，为保证良好的观看体验，视频起播时延宜<1s，视频片段切换时延在采用预加载技术后宜<0.5s，用户拖拽进度条后继续播放的时延宜<2s。

互动类时延可进一步细分为基础点击操作时延和体感操作时延。基础点击操作时延指用户通过手指点击、鼠标点击、滑动屏幕等操作方式产生的时延，主要由互动组件进行响应，基础点击操作的系统响应时延宜 $<50\text{ms}$ 。体感操作时延指通过重力感应、语音识别、人脸表情识别、动作姿态识别等操作方式产生的时延，由于上述操作的响应需依赖硬件、AI 模型或服务端能力，且输入方式本身也需要一定的时间，因此，体感操作的系统响应时延宜 $<2\text{s}$ 。

在实际的互动视频内容制作过程中，除通过技术手段尽可能降低操作响应时延以外，还可通过转场镜头的设计弥补因系统时延给用户观看连贯性带来的影响。

3.7 跨平台兼容实现

跨平台标准是互联网环境下视听网站间建立沟通的基础和互联互通的前提，将引领技术发展和应用创新。建设互动视频跨平台标准体系，既是打造平台技术体系的有力抓手，又是推动平台生态发展的关键环节。

在网络视听行业中有多家视频播放应用，每家又有多个播放平台。为使互动视频可在不同播放平台播放，应考虑互动视频数据格式、播放系统等多个层面的规范统一，如图 3-7。



图 3-7 互动视频跨平台标准化体系

做到跨平台兼容，应建设统一标准规范来消除或适配平台之间的差异。互动视频跨平台运行架构如图 3-8，跨平台建设主要考虑数据协议标准和接口 API 标准两方面的规范统一。

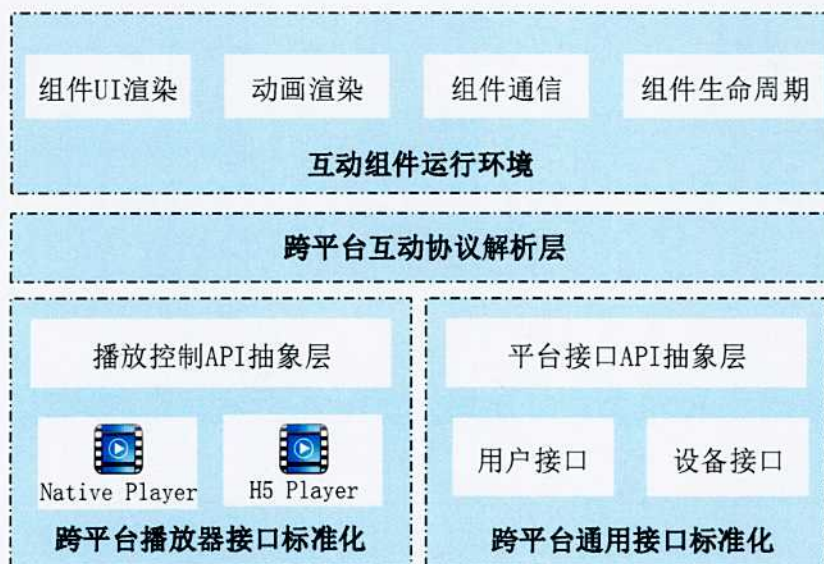


图 3-8 互动视频跨平台运行架构

数据协议标准的规范统一，即应统一不同播放平台应用之间的数据格式标准。对于互动视频数据的组织结构、命名规范、索引脚本等内容，各播放平台需遵循统一的规范约定，使同一部互动视频在不同播放平台的应用间可支持数据协议层面的跨平台兼容。

接口 API 标准的规范统一，即应统一不同终端之间调用的底层 API 接口，统一运行环境。接口分为三部分，播放器控制接口、设备读写接口和用户逻辑接口。各个终端在接口层面的统一，使同一部互动视频在不同终端运行时做到接口调用层面的跨平台兼容。

3.7.1 互动协议跨平台解析标准化

互动视频的数据组织结构包括剧、章、互动节点及播放区间，其数据组织结构如图 3-9。

互动视频由若干章（集）组成，每章（集）互动视频由互动节点和播放区间构成的分支结构组成。互动节点是互动视频内容播放控制的枢纽，其中，播放区间标识 VideoId 字段标识其所属播放区间，互动组件样式模板定义互动界面样式，互动组件参数配置定义标题、倒计时控件、按钮、按钮触发后的播放区间链接等信息。

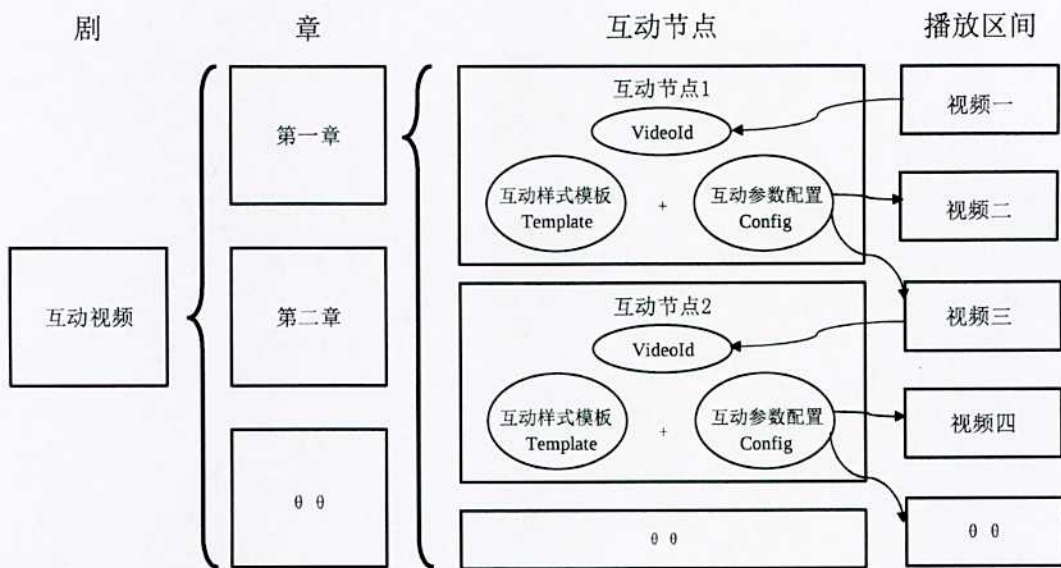


图 3-9 互动视频数据组织结构

从互动视频播放的角度看，互动视频播放器需正确解析这种数据格式，以完成播放逻辑、控制逻辑、用户交互逻辑等功能。在不同播放平台间，应统一互动视频数据格式规范，具体包括以下方面：

- (1) 数据脚本文件集
- (2) 命名规范
- (3) 索引脚本说明文件及定义
- (4) 基础信息、全局信息
- (5) 互动组件样式
- (6) 互动组件参数
- (7) 互动因子

互动视频的主要数据类型见表3-5，包括：剧、章、播放区间、互动节点、互动组件样式模板和互动组件参数配置。通过定义互动视频的主要数据类型，各互动视频播放系统应具备以下两方面的能力：

- 支持自身平台的脚本导出为标准统一的格式；
- 支持将标准统一的脚本文件导入为自身平台支持的格式。

表 3-5 互动视频主要数据类型

字段名	意义	说明
Drama	剧	一部完整剧情的互动视频
Chapter	章	剧的组成部分，通常一部剧含若干章
Video	播放区间	视频片段
InteractNode	互动节点	视频中封装互动组件及播放区间等交互内容的信息集合
InteractTemplate	互动组件样式模板	描述互动组件的模板样式 例如：剧情选择、视角切换、画面互动等
InteractConfig	互动组件参数配置	描述互动组件的参数配置 例如：选项标题、触发的播放区间信息等

3.7.2 互动播放跨平台兼容适配

在互动视频播放系统中，调度互动播放引擎和播放器的指令程序称为互动程序，它运行在互动播放引擎上层的容器中。互动程序通过互动播放引擎封装的 API 控制互动视频的播放、缓冲、渲染等功能，实现与用户的交互响应，如图 3-10。

为使互动程序在不同平台上运行具有良好的兼容性，需统一规范互动视频播放系统 API 接口。

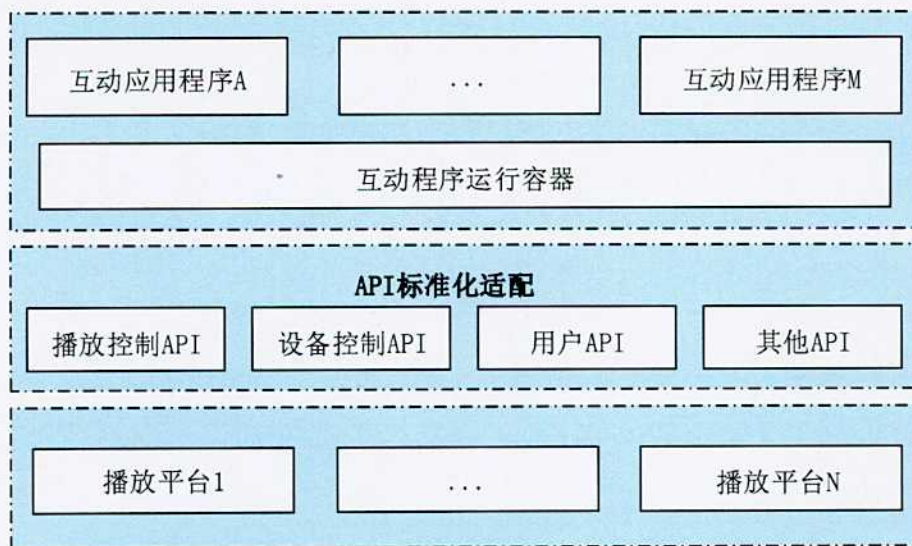


图 3-10 多播放系统间的互动 API 适配

(1) 播放控制接口适配方案

不同平台和终端播放器的能力和接口不同，实现方式也不一致。

a. Android 端使用 Java 语言，对外开放了播放（setPlayerState）、跳转（setSeekTime）等播放器接口；

b. Web 端使用 Javascript 语言，对外开放了播放（play）、跳转（seek）等播放器接口。

互动程序如需在 Native 和 Web 端均可完成互动视频的控制和播放，需要多端研发，成本较高。

另外，由于不同平台提供的播放器能力不一致，或有能力但未对外提供接口，因此，互动程序如需调用播放器倍速能力，则需通过程序来解决不同平台播放器间的能力差异。

通过建立适配层可有效解决不同播放器接口和能力的差异问题，适配层能够屏蔽不同播放平台的差异，统一接口供互动程序调用，并提供相应的能力以满足不同场景需要。跨平台播放器适配层架构如图 3-11。

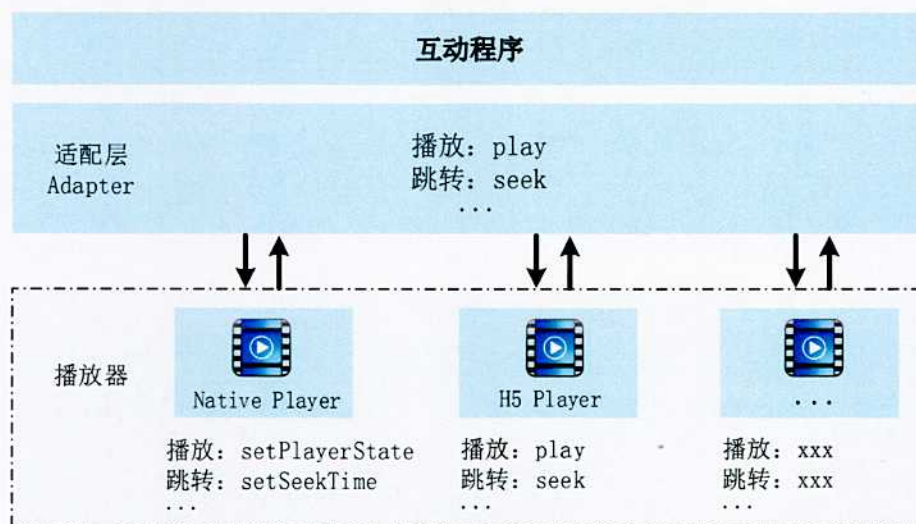


图 3-11 跨平台播放器适配层架构

当互动程序需操作播放器时，将向适配层发送指令，由适配层调度平台的播放器接口实现对播放器的控制，同时会返回不同平台的接口处理结果，并封装成统一的格式回传至互动程序。

(2) 设备控制接口和用户接口适配方案

设备控制接口是播放终端设备的一些传感器向上提供的数据接口，例如：获取屏幕划动轨迹、获取地理位置、获取网络类型、重力感应、读取麦克风、控制摄像头等，这些数据的读写是终端系统提供给播放应用的，播放应用再向互动播放引擎提供接口，互动播放引擎需适配不同播放应用所提供接口的差异，按照统一规范提供 API，供互动程序调用。

用户接口是播放应用向上提供的用户逻辑相关的数据接口，例如：获取登录 ID、读取历史观看记录、数据采集上报、获取头像昵称信息等，这些接口是播放平台向互动播放引擎提供的后台服务接口，互动播放引擎需适配不同播放平台向上提供接口的差异，按照统一规范提供 API，供互动程序调用。

除统一规范 API 以外，还应考虑容错问题。在运行互动逻辑时可能出现各种异常，例如：接口读写未取得用户授权、接口调用超时、用户终端设备异常、播放平台暂不支持某种底层接口等问题。因此，在定义 API 时应充分考虑异常情况，统一错误返回码或返回默认值，以保证互动程序在调用 API 时可根据各种异常做相应处理，达到友好的观影体验。

互动程序通过调用设备控制接口、用户接口以及其他相关接口来完成播放器控制以外的底层调用逻辑，实现丰富的互动形式。

4. 互动视频安全

4.1 内容安全

为营造风清气正的视听环境，针对传播的视频、文字和音频等内容，应从内容质量和封建迷信、恐怖暴力、淫秽色情、侮辱诽谤等违规问题方面进行审核。互动视频相比于传统视频具有更多的内容分支和图文素材，因此，互动视频的审核要求也更复杂，应从内容合规性、合理性等两方面进行内容安全审核。

4.1.1 内容合规性安全审核

内容审核过程中，应检查互动视频故事线的所有分支路径，查看所有的视频片段及互动组件中的图片、文字、音频等信息，确保视频片段内、互动过程中不包含暴力、色情等违规信息。

传统的审核方式，完全依赖于人工审核，存在成本高、速度慢、审核标准不一等弊端。近年来，随着深度学习技术在图像及视频分析领域的显著进展，基于深度学习的自动机审的性能逐步提升。由于互动视频内容审核的工作量较大，因此可采用人工、机器同步审核的方式，提高审核效率。

4.1.2 内容合理性安全审核

内容审核过程中，应对互动操作前后不同视频片段内容衔接的合理性进行审核，确保互动操作前后衔接视频片段的逻辑合理性，不出现隐晦不良信息等。

此外，还应审查互动视频所有分支路径的合理性，避免互动视频播放中出现内容冲突和缺失，确保所有分支路径最终均可走向剧情结局，不出现某一支无限循环播放的情况。

4.2 版权保护

视频盗版通常发生在两个环节：一是在内容制作、报审流程中，由于片源制作、传递、下载、送审等环节的疏漏所致；二是内容上线后被非法录制或技术破解所致。出于对 IP 版权的保护，自上世纪八十年代起，各影视机构就开始在版权保护技术领域持续投资，推动了版权保护相关技术和产业的发展，例如：谷歌的 Widevine、微软的 Playready 和苹果的 Fairplay 等方案，同时中国也制定了

ChinaDRM 标准和测试规范。

此外，近年来区块链技术蓬勃兴起，该技术已逐步在各行业普及应用。作为一种实现可信数据信息记录的数据结构，区块链综合运用非对称密钥、数字摘要等密码学技术，通过多方见证的共识机制和分布式存储方式，实现数据存证记录的可回溯、防篡改、可验证，可为广播电视和网络视听领域的视音频节目等数字化资产提供内容流转、交换、审核、校验等服务，以此降低制作、传播、消费不同领域主体间的信任成本，提升协同效率。区块链提供数字合约功能可进一步与版权管理系统结合，实现版权确权、交易和使用情况统计的智能化，减少人工干预的影响。随着区块链技术的不断演进，利用区块链技术无中心、高可信、可溯源的特点，能够保障视频内容的溯源取证与防篡改，有效提高版权保护环节的可靠性。

由于互动视频对应的互动脚本中包含了大量与视频内容相关的文字、图片、音频等信息，且互动视频相比传统视频具有更多的视频片段和互动节点，因此版权保护的内容更复杂。在互动视频的制作、审核、分发、播放等环节应分别进行相应的版权保护，并通过禁止下载、内容加密、溯源取证等措施保障互动视频的版权安全。

4.2.1 互动视频制作过程版权保护

对于离线制作，制作机构/个人应负责视频的版权保护。对于在线制作，互动视频服务平台应通过技术和管理手段保护视频和素材不被第三方获取，例如通过隐写水印标记上传视频和素材，以区别于制作机构/个人保存的视频和素材，达到界定版权保护的责任目的。借助区块链技术、DRM 技术可甄别视听素材的版权来源，规避版权使用的法律风险。

4.2.2 互动视频审核过程版权保护

在互动视频审核过程中，可对所有分支视频片段和互动视频的完整封装版进行加密，进行版权保护。如发生信息泄露，可通过溯源技术进行盗版用户身份溯源。此外，借助区块链可构建节目内容审核查询体系，建立节目审核的历史信息档案，避免过量重复审核，有效提高审核流程的效率。

4.2.3 互动视频分发过程版权保护

互动视频分发前、分发中可通过 DRM、区块链进行保护。除视频流数据，互动脚本中包含了全部的故事线信息，是互动视频情节发展的重要依据，一旦泄露将会导致用户失去好奇心和观看兴趣。因此互动脚本在分发的过程中应进行加密，防止泄露及篡改，一旦发生内容泄露，可通过区块链的区块前后数据标识逐级关联追溯本源，确定泄露主体。还可通过区块链进行存证，一方面可对分发内容进行校验，防止分发过程中的内容篡改；另一方面，可验证分发任务履约的真实性，从而建立内容提供者和分发服务者间的互信。

4.2.4 互动视频播放过程版权保护

互动视频播放过程中，可采用 DRM 技术进行保护。条件允许时可采用区块链溯源技术进行保护，一旦发生盗版，立即进行用户身份溯源。借助区块链技术还可对视听节目播放量、播放时长等基础统计信息进行可信存证，作为版权使用的证明，并反映内容资产的价值。

4.3 介质安全

视频、图片等互动视频的核心数据存储在服务器的介质中，应考虑数据的容灾容错问题，避免因介质物理损坏或丢失导致视频内容的损失。应建立介质安全管理制度，考虑内容安全问题，该机制宜包括介质管理、容灾管理等方面。

4.3.1 介质管理机制

应对脱机存放的媒资介质进行控制和保护，以防止被盗、被毁、被修改、信息泄露等事件发生，宜从以下三方面进行介质管理。

(1) 介质访问安全：仅授权人员可以访问介质，对重要介质中的数据和软件采取加密存储；

(2) 介质存储安全：磁盘存储在专用的磁盘保存柜中，保持适宜温度，防止潮湿、强磁场等；

(3) 介质使用安全：需携带介质外出时应经过授权，并保障此期间的介质安全。介质被送出维修或销毁前应先清理介质中的敏感数据。此外，对存放在介质库中的介质应定期进行完整性和可用性检查，确认其数据或软件未受到损坏或丢失。

4.3.2 灾备管理机制

(1) 异地备份，在保证数据加密不泄露的前提下，可采用异地多实例部署、多物理磁盘备份存储等方法进行容灾管理；

(2) 网络冗余，通过使用不同运营商的带宽服务，保证在某个运营商网络故障时，网络传输不受影响。对于单个运营商带宽，应不低于系统运行峰值所需带宽的 1.5 倍；

(3) 计算资源动态调配，使用云技术，在业务高峰期可调度计算资源保障核心业务；

(4) 动力备份，机房动力系统应有多种应急措施，包括多路稳定动力电路、UPS、燃油发电机等，并安全储备充足燃料。

4.4 运行环境安全

4.4.1 互动脚本安全

开发者编写互动脚本的语言应确保安全可靠，脚本可进行充分封装，且无法通过互动视频服务平台及播放系统以外的手段进行访问。此外，互动脚本应保证故事线互动空间的封闭性，禁止通过脚本链接至外部未经授权地址的行为。终端在互动视频播放过程中会在本地存储器内暂存互动脚本，需确保在此过程中脚本始终以加密形态保存。

4.4.2 隔离与隐私保护

为确保互动视频间的逻辑隔离及用户账号的隐私保护，互动视频系统可对互动视频内容与播放器端的容器环境进行域名隔离。此外，用户账号信息可进行加密或非明文转换，创作者或其他第三方无法直接访问指定用户的隐私信息。

4.4.3 沙箱执行环境

由于互动组件在播放终端执行，为保障互动视频播放系统的安全，需要在运行时通过沙箱隔离互动引擎和互动组件，确保执行环境安全。

4.5 平台与传输安全

4.5.1 平台安全

平台自身的安全是动态的，应从各方面详尽分析，选择安全性尽可能高的操作系统和硬件平台，并对平台进行安全配置。此外，应加强登录过程的认证，对用户访问网络资源的权限进行严格的认证和控制，对口令进行加密、更新和鉴别，并设置用户访问目录和文件的权限等，同时，应加强对网络关键设备的保护，制定严格的网络安全规章制度，通过技术和制度等方面的把控，提高平台的安全性。

4.5.2 传输安全

传输安全可分为三个部分：数据传输安全、报文传输安全和传输链路安全。在数据传输安全方面，互动视频服务平台提供的接口需要保证各种交互数据安全传输。在报文传输安全方面，由加验签和加解密保证报文安全控制，所有报文均需加验签，涉及敏感信息的报文应做加解密。在传输链路安全方面，可采用加密认证、文件伪装、限制引用页等技术方案，结合 HTTPS 传输通道及 TLS 1.2 协议，通过单向认证的方式，避免数据在网络传输过程中被截取或窃听。

4.6 用户隐私安全

用户在观看互动视频内容、体验互动视频服务时，播放系统可能会收集、存储并使用用户的相关信息，包括观看设备信息、观看账户信息、互动操作信息等，从用户角度，应建立用户隐私授权机制和用户数据防篡改机制，充分保障用户隐私安全。

4.6.1 用户隐私授权机制

授权机制，是手机、平板等终端操作系统安全机制中的一部分，互动视频在播放过程中如调用相机、麦克风、相册、通讯录、网络、位置、通知等涉及用户隐私的应用程序，必须经过用户授权，应采用授权弹窗等形式通知用户，由用户选择允许或拒绝，如用户允许，则授权通过后，方可调用相关功能或程序；如用户拒绝，则表明相关功能或程序不允许使用，用户在进入相应页面时，播放系统应给予提示说明。

4.6.2 用户数据防护机制

用户在观看互动视频内容、体验互动视频服务时，所产生的观看数据、行为数据、消费数据、互动数据等均属于用户隐私的一部分，各平台应结合现阶段技术系统现状，从三个方面做好防护：其一，建立用户防泄露机制，在用户数据存储、传输和使用过程中，及时发现并识别用户数据隐患，防止用户数据泄露；其二，建立用户数据防篡改机制，通过数据加密等技术方案，避免通过客户端直接修改用户数据，以防用户数据被恶意篡改；其三，建立用户数据防丢失机制，在人为原因、自然原因、软硬件原因等各类导致数据丢失的事件发生前，通过数据备份、缓存等机制，防止用户资产和数据遭受损失。

5. 互动视频应用展望

2019年8月，“中国广电·青岛5G高新视频实验园区项目”战略合作备忘录签署及揭牌仪式在青岛举行，中宣部副部长、国家广播电视总局局长、党组书记聂辰席在会上指出：“在5G环境下推动高新视频新供给，引领高新视频新业态，拉动高新视频新消费，为满足人民群众新期待注入强大动能，为国家数字经济发展和数字中国建设作出新的更大贡献”。加快高新视频服务模式与关键技术的探索与应用，是当前面临的重大课题和关键任务。互动视频作为高新视频业态的重要组成部分，应加速推动互动视频在内容生产、传播方式、业务形态等多方面的发展。

随着5G传输网络的覆盖普及，5G网络所特有的“大带宽、低时延、广连接”的优良特性为互动视频发展应用创造了条件。未来，互动视频也可应用于VR视频制播领域，打造360度互动型VR视频，并更广泛地应用于影视、教育、医疗、文化、旅游、体育、电商等各行业。

希望在本白皮书的引导下，各级广电机构、行业内外企业共同努力，积极推进互动视频的应用发展，丰富视听形态，为广大人民群众带来更美的视听体验，并通过互动视频引导广大人民群众在内容建设过程中表达观影需求、深度参与互动，进一步增强广播电视内容核心竞争力，推动广播电视和网络视听行业高质量发展创新发展。

附录 A 互动影视剧案例

A.1 案例 1

A.1.1 案例概述

2019年6月20日,爱奇艺上线互动影视剧《他的微笑》。作为爱奇艺首个落地的互动影视作品,该剧充分运用了互动视频中的剧情选择互动能力,突破了传统影视叙事模式,为用户观影打开“自由之门”、“上帝视角”。全剧共设置21个选择节点、17种结局,所有剧情总时长约200分钟,最短故事线仅5-10分钟,最长达50-60分钟。

A.1.2 案例详情

《他的微笑》通过向用户开放剧情走向的选择权,让用户代入并沉浸剧情角色,根据个人喜好做出选择,展开与之相关的故事情节,从“观看者”变身“参与者”实现沉浸式的观影体验。在故事中,用户会跟随女主角共同做出选择,而每一个选择都会对结局产生重大影响,使用户更能沉浸于剧情中。具体来说,《他的微笑》的互动能力表现在如下方面。

(1) 剧情选择能力

在诸多关键剧情节点,用户可以基于自身立场或偏好进行选择,参与、影响剧情发展,进入到不同的剧情内容进行体验,如图A-1。



图 A-1 互动影视剧《他的微笑》互动能力

(2) 故事线结构

故事线记录用户在互动视频中的观影路径和每一个选择，并可满足用户从已观看的任意一段剧情、或一个互动选择进入继续观看，方便用户回溯剧情、重新探索多样化的剧情，如图 A-2、A-3。

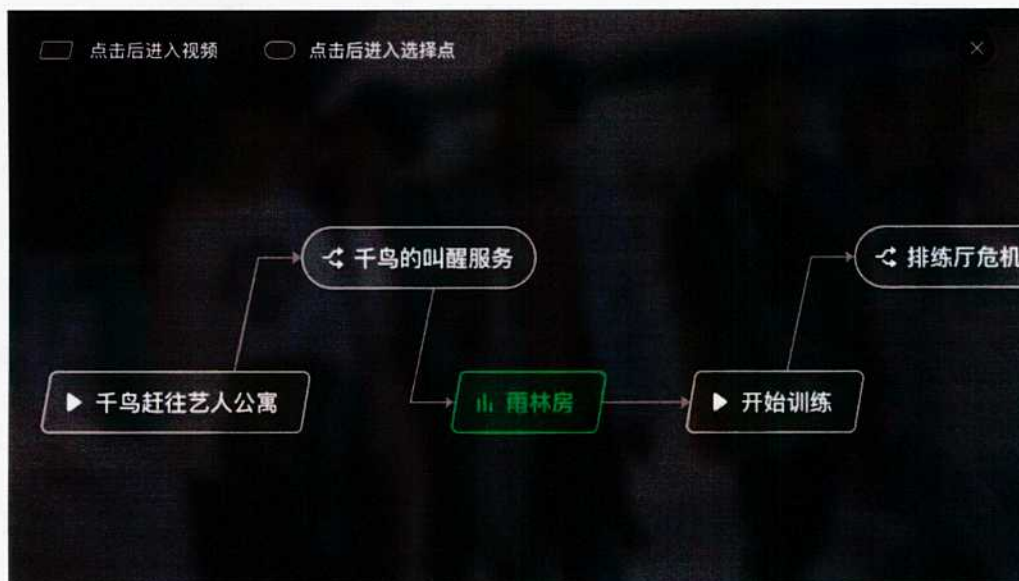


图 A-2 互动影视剧《他的微笑》故事线结构

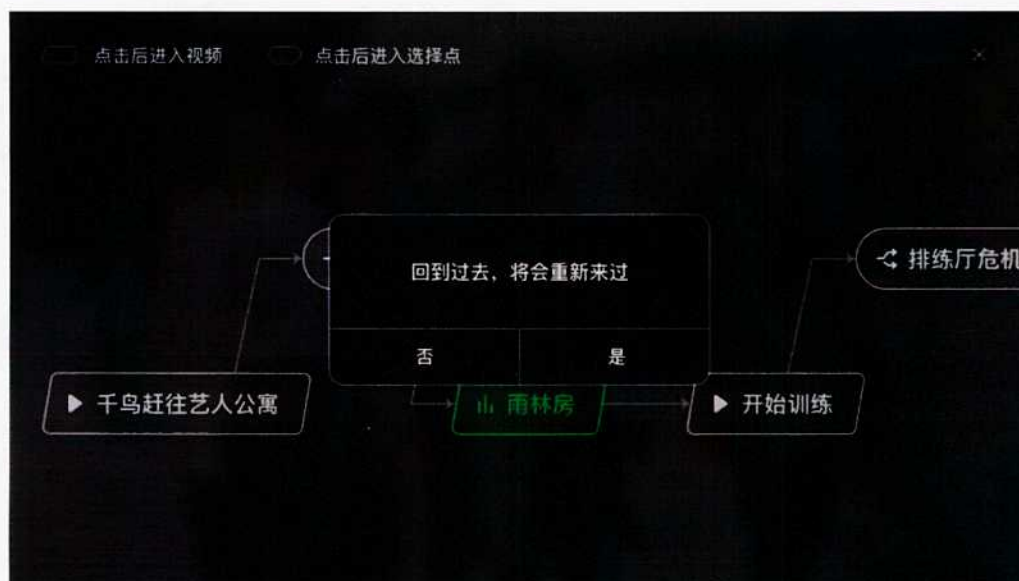


图 A-3 互动影视剧《他的微笑》故事线能力

A. 1.3 案例效果

对用户来说，互动视频革新内容表达与用户的连接方式，以互动视频为代表的互动沉浸体验将全面更新用户的娱乐方式。在有效观看中，约有 60%的用户主

动参与互动，其中约 35% 的用户达成多个结局。在参与互动的用户中，人均互动 9.5 次，人均观看 3.5 个结局。

对爱奇艺而言，《他的微笑》促使互动视频实验产业化运作落地，有助于在内容制作、播出机构间逐步构建一个面向互动视频制作、发布及运营的生态循环。

A.2 案例 2

A.2.1 案例概述

2019 年 9 月 15 日，由腾讯视频和动视无界联合出品的全屏互动影视剧《因迈思乐园》在腾讯视频上线，该剧共由 5 集组成，其单集单支线时长为 5-10 分钟，全剧共设置 30 余个互动节点、3 种结局，通过“全屏互动”形式，引入了 6 种不同形态的互动方式。

A.2.2 案例详情

《因迈思乐园》在故事过程中采用了多种新颖的互动形式，具体来说，《因迈思乐园》的互动能力表现在如下方面。

(1) 剧情选择能力

剧情选择一般通过设计关键剧情并增加倒计时，让用户在观影过程中快速反应做出选择，并由此进入不同的分支剧情中。

在《因迈思乐园》中，运用了大量限时剧情选择，用户在很多时候需及时在较少的时间内做出选择，利用互动的形式放大了原本惊险内容中的紧张刺激感，如图 A-4。



图 A-4 互动影视剧《因迈思乐园》互动能力 1

（2）视角切换能力

《因迈思乐园》在作品第一章中同步引入视角切换组件样式，用户以女主角视角在开车过程中，可随时点击切换“查看路况”，宛如导演一般灵活切换视角镜头，给用户互动带来很好的互动引导及操作权力感，如图 A-5。



图 A-5 互动影视剧《因迈思乐园》互动能力 2

（3）画面互动能力

该剧中借助画面互动能力，要求用户在极短的时间快速做出选择和反应，一般通过设置较短时间的倒计时，适合在惊险或动作类题材的作品内容中使用。《因迈思乐园》主角在寻找线索中经历多种紧张惊险的时刻，在较短的时间要求用户快速做出反应，引导角色成功脱险，如图 A-6。



图 A-6 互动影视剧《因迈思乐园》互动能力 3

（4）VR 全景互动能力

通过 VR 全景技术，结合限时设置增加趣味度和紧张感，在寻找线索时，用户可选择不同的监控视角观看。在数据分析技术的支持下，剧中的每一个选择都成为变量和“命运伏笔”，影响主角最终回到现实的几率，形成了巨大的戏剧张力，如图 A-7。



图 A-7 互动影视剧《因迈思乐园》互动能力 4

(5) 故事线结构

《因迈思乐园》在作品中专门设计了故事线，能够让用户清楚地看到自己选择了哪一条故事支线，可再次返回何处节点以体验另一种可能，使得故事结构清晰明了，素材得以充分利用，如图 A-8。



图 A-8 互动影视剧《因迈思乐园》故事线结构

(6) 互动小游戏能力

《因迈思乐园》作品中引入了小游戏互动方式，用户可以代替女主角实际参与小游戏，游戏是否成功通关将直接影响剧情的走向，如图 A-9。

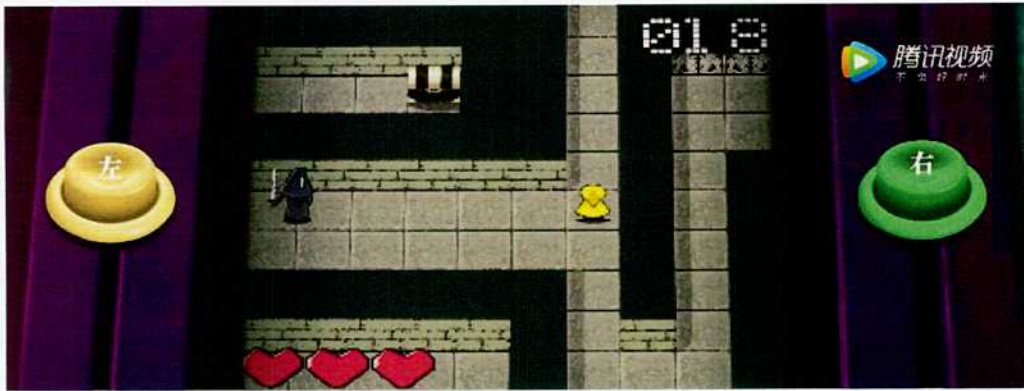


图 A-9 互动影视剧《因迈思乐园》互动能力 5

A. 2.3 案例效果

《因迈思乐园》作品中引入了丰富的互动玩法，包括全屏互动（无需点击选项文字）、限时互动（默认选择左侧选项）等，此外，通过设置富有代入感的“冲过去”单选项、“抓住栏杆”动作选项、剧情小游戏，极大增强了观影娱乐感。

《因迈思乐园》上线不到 3 小时，播放量达 500 万，截至 2020 年 6 月 1 日，播放量已超 1700 万，其中，百余万用户参与了内容互动体验，人均互动 6.5 次，人均观看 2.8 个结局，80% 的用户参与了互动体验，65.8% 的用户会反复体验多个互动节点的不同选项，35.9% 的用户观看了全部故事线内容。

A. 3 案例 3

A. 3.1 案例概述

2020 年 2 月 29 日，优酷上线平台首部体感互动剧《娜娜的一天》，该剧共由 6 集组成，单集长度为 10 分钟，全剧共设置 200 个互动节点、22 个结局，剧中引入了语音识别、人脸识别、触感点击等多项体感互动方式。

A. 3.2 案例详情

《娜娜的一天》创新的利用第一人称视角的叙事方式，且兼容点击选项互动和语音聊天互动，运用了语音识别、人脸识别等体感互动方式。

(1) 剧情选择能力

用户观看至互动节点时出现分支选项，并开始计时，如在倒计时结束前用户选择某一分支，则进入对应情节；否则进入默认选项。观众可基于自身立场或偏好进行选择，并获得成就与满足感，如图 A-10。



图 A-10 互动影视剧《娜娜的一天》互动能力 1

(2) 画面互动能力

剧中依据人物角色引导和画面视觉提示，引导用户在指定时间内，通过触屏方式完成指定动作（例如：擦拭刮奖、画爱心等），并根据用户完成度决定剧情发展脉络，如图 A-11。



图 A-11 互动影视剧《娜娜的一天》互动能力 2

(3) 语音识别方式

在《娜娜的一天》作品中的内容关键节点，大量使用了语音识别互动方式，在用户授权 APP 可调用麦克风的情况下，根据剧中视频内容的引导，在指定时间内引导用户开口与主角对话，根据用户话语内容决定剧情走向，营造剧中主角和用户面对面聊天的场景，增强用户的身临其境感，如图 A-12。



图 A-12 互动影视剧《娜娜的一天》互动方式 1

(4) 人脸识别方式

在用户授权 APP 可调用摄像头情况下，根据剧中角色的话语引导和播放提示，在手机屏幕前做出指定脸部动作（例如：脸部左右转动），并判断动作是否符合要求（例如：要求脸部左转时，判断是否为左转、是否达到指定度数等），进而决定后续内容的发展和走向，如图 A-13。



图 A-13 互动影视剧《娜娜的一天》互动方式 2

A. 3.3 案例效果

《娜娜的一天》作品中创新应用了语音识别、人脸识别等体感互动方式，使用户的参与感和沉浸度均有所提升，从用户互动频率来看，体感类互动方式与视频内容的结合有效提高了用户互动频次，超过 90% 的用户参与了剧中互动环节，人均互动次数超过 10 次，其中大多用户乐于使用语音聊天的互动方式，人均复看次数超过 3 次，与点播场景相比，互动形式的内容更容易吸引用户留存；从受众群体角度看，该类型的互动剧更受年轻群体喜爱，其中 18-30 岁男性用户观众占比超过 40%。

附录 B 互动综艺案例

B.1 案例概述

互动能力在综艺内容上可以有丰富的应用。例如，通过视角切换，内容创作者可以在综艺节目中同时为用户提供多个视角的内容，对应到候场区、主舞台等不同的场景。此外，通过画面交互，内容创作者可以极大地丰富节目内容，将剧情主线外有趣、有料的内容在视频内提供给用户自行探索发现。

爱奇艺综艺节目《中国新说唱》2019 中综合运用了多项互动能力，在节目中运用剧情选择能力，为用户带来了充满参与感的综艺体验，并通过互联网互动广告，使用户根据喜好自主选择广告剧情，为互联网视频广告带来了一定的创新。

B.2 案例详情

(1) 剧情选择能力

在《中国新说唱》第十一期中“明星帮唱环节”部分，选手在面临嘉宾合作舞台的双向选择时，两人画面分别弹出两个互动按钮供用户选择，以此来观看不同的分支剧情，如图 B-1。



图 B-1 互动综艺《中国新说唱》互动能力

(2) 互联网互动广告

在《中国新说唱》2019 第八期节目中，爱奇艺上线了互动视频广告，将互动技术运用到娱乐广告层面。该广告将综艺 IP、原生创意广告与互动相结合，依托于《中国新说唱》2019 节目内容，邀请节目人气选手以 rap 演唱形式将

广告方的产品卖点和品牌内核完美呈现，临近尾声，画面弹出的两个互动按钮，用户可以通过选择不同按钮来观看不同分支剧情，如图 B-2。



图 B-2 互动综艺《中国新说唱》互动应用

B.3 案例效果

《中国新说唱》2019 是通过使用剧情选择互动能力为用户带来了新颖的综艺内容体验。节目中通过引入互动视频广告，使总曝光次数得到了有效提升，原始广告播放次数提升约 18%，21% 的用户观看了多个分支内容，加深了用户的品牌认知度和好感度，为视频广告创新带来了新思路。

附录 C 互动短视频案例

C.1 案例概述

2019年9月，腾讯微视和北京急救中心联合推出《急救小星君》账号，主推急救知识科普类互动短视频内容，目前已累积发布30余个互动短视频，平均视频时长不超过1分钟，主要借助剧情选择互动能力使用户参与其中。

C.2 案例详情

每集视频内容会模拟一个常见的急救案例，并在关键时刻通过倒计时增加紧张感，促使用户根据个人知识积累快速做出选择，用户做出选择后，视频会给予正误反馈，回答错误的用户将看到剧情危机爆发，并通过视频内专业的讲解学习正确的急救知识，回答正确的用户则看到危机顺利解除，并掌握相应的急救原理，如图C-1。



图 C-1 互动短视频互动能力

C.3 案例效果

科普类内容结合互动短视频，可通过对错反馈等交互方式，加深用户对科普知识的印象。相较于传统视频动画式、说教式的演示，优势较为突出。《急救小星君》让用户直面“生死”选择，带来更沉浸的体验，目前，账号关注率、互动率皆高于知识品类下其他普通账号均值。

参考文献

- [1] ETSI TS 201 812: 2003 Multimedia Home Platform (MHP) Specification
1.0.3
- [2] HbbTV: 2018 HbbTV 2.0.2 Specification
- [3] ETSI TS 102 796: 2016 Hybrid Broadcast Broadband TV
- [4] IPTV FJ STD-0010: 2013 Integrated broadcast-broadband system
specification
- [5] ATSC A/344: 2019 ATSC 3.0 Interactive Content
- [6] ITU-R Rep. BT.2267-10: 2019 Integrated broadcast-broadband systems